From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Mathematics

2022

Gibbs measures on sparse random graphs

Airey, Dylan

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01v118rh71p

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Gibbs measures on sparse random graphs

Dylan Airey

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mathematics
Adviser: Allan Murray Sly

May 2022

© Copyright by Dylan Airey, 2022.

All Rights Reserved

Abstract

Many combinatorial optimization problems can be phrased in terms of Gibbs measures on labellings of the vertices of a graph. Techniques and insights from statistical physics have been applied in recent years to find the solvability thresholds as well as the structure of solutions for many of these problems. When the underlying graph is random, two key quantities of interest are the quenched and annealed free energy densities. The annealed free energy density is much easier to compute, and for some models, like the Ising ferromagnet, the quenched and annealed free energy densities agree at all temperatures. For others, the system undergoes a phase transition, and the two disagree at low temperature. We show this latter phenomenon is typical in the sense that for an open dense subset of interaction functions, on a random regular graph the quenched and annealed free energy densities of the associated Gibbs measure disagree for sufficiently low temperatures.

We also present joint work with L. Bowen and F. Lin in which we consider a particular model, 2-colorings of random regular hypergraphs, to answer an open problem in sofic entropy theory. A sofic approximation to a countable group is a sequence of partial actions on finite sets that asymptotically approximates the action of the group on itself by left-translations. A group is sofic if it admits a sofic approximation. Sofic entropy theory is a generalization of classical entropy theory in dynamics to actions by sofic groups. However, the sofic entropy of an action may depend on a choice of sofic approximation. All previously known examples showing this dependence rely on degenerate behavior. We exhibit an explicit example of a mixing subshift of finite type with two different positive sofic entropies.

Acknowledgements

I would like to thank my advisor Allan Sly for his guidance throughout my years in graduate school. I am very lucky to have had the opportunity to work so closely with such a great mathematician, as well as an exceptional advisor. I also cannot thank Lewis Bowen enough for his mentorship as an undergraduate and graduate student.

Thank you to my other coauthors Stephen Jackson, Dominik Kwietniak, Frank Lin, Bill Mance, and Roman Nikiforov. It has been a joy to do mathematics with you all, and thank you for your patience during the pandemic. Thank you to Evita Nestoridi and Ramon van Handel for serving on my FPO committee and to Michael Aizenman for being a reader for my dissertation. I greatly appreciate your helpful feedback and very interesting questions.

Thank you to Dor Elboim, Shuangping Li, Danny Nam, and Lingfu Zhang for making being a graduate student in probability such a great experience. Thank you also to all of the other graduate students in the department, there are too many of you to name, but each of you made Princeton a very special place to be.

I am greatly appreciative of the National Science Foundation for funding my research with the Graduate Research Fellowship, as well as the Princeton mathematics department for their funding and the freedom they provided to pursue my research.

Thank you to my friends and family, you have given me strength during the many difficulties of grad school, and I am extremely lucky to have you in my life.

To my wife Arlenne and my parents Alan and Lynda.

Contents

Abstract iii

Acknowledgements iv

1 Introduction **1**

1.1 Statistical physics models on graphs 2

1.1.1 Gibbs measures 2

1.1.2 Random graph models 3

1.2 Gibbs measures on random regular graphs and replica symmetry . . . 4

1.2.1 Related works 8

1.3 Sofic entropy 9

1.3.1 Random sofic approximations 11

2 The quenched and annealed free energy for generic interactions **13**

2.1 Non-backtracking matrix of regular graphs 13

2.2 Relationship between ground state energy and $\mathcal{RF}$ 15

2.3 $\mathcal{G}$ is open 17

2.4 Construction of interactions in $\mathcal{RF}$ 17

2.4.1 Bounding $\mathbb{E}[\min H_{\Phi+\epsilon\Psi}]$ 19

2.4.2 Bounding the probability of $\min H_{\Phi+\epsilon\Psi} < \mathbb{E}[\min H_{\Phi+\epsilon\Psi}] - \delta n$. 23

2.5 Subshifts 27

2.5.1 Lack of valid microstates 29

2.5.2 Dobrushin mixing 31
2.6 Construction of interactions in $\mathcal{RF}$ for subshifts 32

3 Dependence of topological sofic entropy on the choice of sofic approximation 36

3.0.1 Proper colorings of random hyper-graphs from a statistical physics viewpoint 36
3.0.2 The action 40
3.0.3 Sofic entropy of the shift action on proper colorings 42
3.0.4 Random hyper-graph models 42
3.0.5 The strategy and a key lemma 44
3.1 Topological sofic entropy 46
3.2 Reduction to the key lemma 48
3.3 The first moment 54
3.3.1 Almost proper 2-colorings 55
3.3.2 Equitable colorings 61
3.4 The second moment 64
3.4.1 Analysis of ψ_0 and the proof of Lemma 3.0.6 inequality (3.2) . 76
3.4.2 Reducing Lemma 3.0.6 inequality (3.3) to estimating the local cluster 86
3.5 The local cluster 90
3.6 A Markov process on the Cayley hyper-tree 94
3.6.1 Local convergence 95
3.6.2 The density of the rigid set 107
3.7 Rigid vertices 114

Appendices 123

A Topological sofic entropy notions 124

B Concentration for the planted model **128**

B.0.1 General considerations 129

B.0.2 Specific considerations 134

Chapter 1

Introduction

Applying techniques from statistical mechanics to combinatorial optimization problems on graphs has been a very active and fruitful area of research. Such methods have been used to make sweeping predictions about the optimal values of such problems, as well as the structure of the space of solutions. In many cases, these predictions have been rigorously verified, such as in the chromatic number of random regular graphs [1], the maximum size of independent sets on random regular graphs [2], the size of both the Max-Cut and maximum bisection of random regular graphs and sparse Erdös-Rényi graphs [3], random k-SAT [4], and community detection in the stochastic block model [5].

Each of these models has required a careful analysis and difficult ad-hoc combinatorial estimates which are generally applicable only to that model. One of the aims of this dissertation is to study the behavior of a *generic* model.

The other aim is to use a specific model, 2-colorings of random regular hypergraphs, to answer an open question in sofic entropy.

1.1 Statistical physics models on graphs

1.1.1 Gibbs measures

Given a finite alphabet $\mathcal{A}$, we may consider a collection of n variables taking values in $\mathcal{A}$. We call the set $\mathcal{A}^n$ the space of **spin configurations**. In analogy with physics, we think of these variables as the spins of a collection of interacting particles. These interactions give rise to an energy function, the **hamiltonian**, which for our purposes will be a function $H : \mathcal{A}^n \to \mathbb{R}$. Associated to each hamiltonian H is the **Gibbs measure** at inverse temperature β. This is a probability measure $\mu_H \in \text{Prob}(V^{\mathcal{A}})$ given by

$$\mu_H(\chi) = \frac{1}{Z} e^{-\beta H(\chi)}$$

for all $\chi \in \mathcal{A}^n$ where Z is a normalizing constant called the **partition function**. According to physics, a closed system (where total energy is constant) will tend toward an equilibrium which maximizes entropy, and the Gibbs measure (or Boltzmann distribution), is precisely this equilibrium. The parameter β balances between energy and entropy. When $\beta = 0$ (infinite temperature), μ is simply the uniform measure over $\mathcal{A}^V$, and entropy is maximized. When $\beta \to \infty$ (zero temperature), μ is supported exactly on the **ground states** of H, those spin configurations with minimum energy.

Of central importance is the free energy density, given by the normalized logarithm of the partition function

$$\varphi(\beta) := \lim_{n\to\infty} \frac{1}{n} \log Z(\beta\Phi).$$

Many quantities of interest can be expressed in terms of $\varphi(\beta)$: for example the average energy density $\lim_{n\to\infty} \langle \frac{1}{n} H \rangle_\mu$ is given by $\frac{\partial}{\partial\beta}\varphi(\beta)$, and phase transitions occur at points of non-analyticity of $\varphi(\beta)$.

A common class of hamiltonians are given by pairwise interactions between particles. These pairwise interactions form the structure of a graph $G = (V, E)$, where

V is our collection of variables.

Example 1.1.1. *Let $V = \{-n, -n+1, \cdots, n-1, n\}^d$ with $\{\mathbf{n}, \mathbf{n}'\} \in E$ when $\|\mathbf{n} - \mathbf{n}'\|_1 = 1$. That is, $G = (V, E)$ is the d-dimensional integer lattice restricted to the hypercube of side length $2n+1$. Let $\mathcal{A} = \{\pm 1\}$, and*

$$H(\chi) = -\sum_{(\mathbf{n},\mathbf{n}')\in E} \chi_{\mathbf{n}}\chi_{\mathbf{n}'}.$$

*This is the **ferromagnetic Ising model** on the lattice, a model for magnetism and one of the simplest statistical mechanics models exhibiting a phase transition (when $d \geq 2$).*

Example 1.1.2. *Let $G = (V, E)$ be the complete graph on n vertices, $\mathcal{A} = \{\pm 1\}$ and for each $1 \leq i, j \leq n$ let g_{ij} be iid standard Gaussian random variables. Then*

$$H(\chi) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{n} g_{ij}\chi_i\chi_j$$

*is the hamiltonian for the **Sherrington-Kirkpatrick model**.*

Unlike the Ising model on the lattice, the Sherrington-Kirkpatrick model has a random hamiltonian. One may then ask for the behavior of the associated Gibbs measure for a typical sample of the hamiltonian. We may consider classes of random Gibbs measures in another way: fixing an interacting term as in the Ising model, but allowing the graph of pairwise interactions to be a random graph.

1.1.2 Random graph models

Throughout this work we will work with random regular graphs and hyper-graphs. There are a number of possible probability distributions one may consider. First, for $n, d \in \mathbb{N}$ the number of d-regular graphs on n-vertices is a finite set, and so one may consider the uniform distribution over this set. This is the **uniform model** for

d-regular random graphs. While conceptually simple, this model is difficult to work with in practice.

Another common choice is the **configuration model**. Here one views $[n]\times[d]$ as a collection of half-edges. Choosing a perfect matching uniformly at random and gluing together these half edges accordingly then yields a d-regular multi-graph. Necessarily nd must be even in order for such a matching to exist.

In this work we will instead work with yet another model, the **permutation model**. We sample d permutations $\pi_{i,n} \in \mathrm{Sym}(n)$ for $1 \leq i \leq d$ independently and uniformly at random. We then consider the multi-graph with $V = [n]$ and $E = \{\{v, \pi_{i,n}(v)\} : v \in V, 1 \leq i \leq d\}$. The graphs generated will be $2d$-regular rather than d-regular.

Note that the configuration model is equivalent to choosing an element of $\mathrm{Hom}(*_{i=1}^{d}\mathbb{Z}/2\mathbb{Z}, \mathrm{Sym}(n)$ uniformly at random where $*_{i=1}^{d}\mathbb{Z}/2\mathbb{Z}$ is the d-fold free product of $\mathbb{Z}/2\mathbb{Z}$, while the permutation model is equivalent to choosing an element of $\mathrm{Hom}(\mathbb{F}_d, \mathrm{Sym}(n))$ uniformly at random.

1.2 Gibbs measures on random regular graphs and replica symmetry

A key property of random regular graphs is that they are **locally tree-like**: for any fixed $R \in \mathbb{N}$, as $n \to \infty$ w.h.p. the radius R neighborhood of an arbitrarily close to 1 fraction of vertices will contain no cycles. This motivates the following construction of Gibbs measures on random regular graphs.

Let $\Gamma = \mathbb{F}_d = \langle s_1, \cdots, s_d \rangle$ be the free group on d generators. Let $G_n = (V_n, E_n)$ be given by the random permutation model, and let $\sigma_n : \Gamma \to \mathrm{Sym}(n)$ be the associated homomorphism. For $v \in V_n$ define $\Pi_v^{\sigma_n} : \mathcal{A}^{V_n} \to \mathcal{A}^{\Gamma}$ by $\Pi_v^{\sigma_n}(\chi)(\gamma) = \chi(\sigma_n(\gamma)(v))$. In words, $\Pi_v^{\sigma_n}$ is the lift to the universal cover Γ of the coloring χ with basepoint v. See

3.1 for further motivation of this mapping.

Definition 1.2.1. *An **interaction function** is a continuous function $\Phi : \mathcal{A}^{\Gamma} \to \mathbb{R}$. The **hamiltonian** $H_{\Phi,n} : \mathcal{A}^{V_n} \to \mathbb{R}$ associated to an interaction function Φ is given by*

$$H_{\Phi,n}(\chi)(\sigma_n) := \sum_{v \in V_n} \Phi(\Pi_v^{\sigma_n}(\chi)).$$

*For an interaction function Φ the **Gibbs measure at inverse temperature** β is the probability measure $\mu_{\beta\Phi,n} \in \mathrm{Prob}(\mathcal{A}^{V_n})$ given by*

$$\mu_{\beta\Phi,n}(\chi) = \frac{1}{Z_n(\beta\Phi)} e^{-\beta H_{\Phi}(\chi)}$$

*where $Z = \sum_{\chi \in \mathcal{A}^{V_n}} e^{-\beta H_{\Phi}(\chi)}$ is a normalizing constant called the **partition function**.*

We suppress the dependence on n and σ_n when it is clear from context.

Example 1.2.2. *The following are interaction functions associated to commonly studied systems.*

- $\mathcal{A} = \{\pm 1\}$ *and* $\Phi_{Ising}(x) = -\sum_{i=1}^{d} x_{1_\Gamma} x_{s_i} + x_{1_\Gamma} x_{s_i^{-1}}$ *is the **ferromagnetic Ising model***
- $\mathcal{A} = \{\pm 1\}$ *and* $\Phi_{af-Ising}(x) = \sum_{i=1}^{d} x_{1_\Gamma} x_{s_i} + x_{1_\Gamma} x_{s_i^{-1}}$ *is the **anti-ferromagnetic Ising model***
- $\mathcal{A} = \{0, 1\}$ *and* $\Phi_{hc}(x) = -x_{1_\Gamma} + \sum_{i=1}^{d} c(x_{1_\Gamma}, x_{s_i}) + c(x_{1_\Gamma}, x_{s_i^{-1}})$ *where* $c(0,0) = c(0,1) = c(1,0) = 0$ *and* $c(1,1) = \infty$ *is the **hard-core model**.*
- $|\mathcal{A}| > 2$ *and* $\Phi_{Potts}(x) = -\sum_{i=1}^{d} \mathbb{1}_{x_{1_\Gamma} = x_{s_i}} + \mathbb{1}_{x_{1_\Gamma} = x_{s_i^{-1}}}$ *is the **ferromagnetic Potts model***

Technically Φ_{hc} is not an interaction function according to our definition since it takes values of ∞. We may work with such interactions by modifying their domain. We consider such cases in 2.5.

Since we are working with random graphs, the hamiltonian and Gibbs measure will themselves be random. While the free energy $\varphi(\beta)$ is a random variable, standard concentration results (see Appendix B) imply that $\log Z(\beta\Phi)$ concentrates about its expectation, and so one may instead analyze

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log Z(\beta\phi)],$$

the **quenched free energy density**. In general this limit may not exist, and even when it does, it is usually extremely difficult to compute. However, by Jensen's inequality one may find an upper bound given by the **annealed free energy density** [1]

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[Z(\beta\Phi)].$$

The annealed free energy density is much more tractable, though still difficult to compute generally.

For many interaction functions, at sufficiently high temperatures the decay of correlations in $\mu_{\beta\Phi}$ leads to the phenomenon of **replica symmetry**. One consequence of this property is that the quenched and annealed free energy densities agree. For d-regular random graphs, the annealed free energy is equal to the replica symmetric prediction of the free energy (see Proposition 14.8 in [6]).

G. Parisi in [7] introduced a scheme, now called the Parisi ansatz, to predict the free energy of the Sherrington-Kirkpatrick model based on the idea of sampling spin configurations from the Gibbs measure independently (these are the replicas which give replica symmetry its name), and considering the distribution of the overlap, the fraction of vertices where the two configurations agree. He predicted a particular

[1] The terms quenched and annealed are taken from alloy production. One quenches an alloy by plunging the hot metal into a cold water bath and freezing the atoms in place, while one anneals an alloy by letting the metal to come to room temperature gradually. This alloys the atoms to move as the alloy cools, the dynamics of the magnetism and atom placement take place over the same time scale.

form for this overlap distribution which implies an interesting clustering phenomenon for the Gibbs measure. The mass of the Gibbs measure is concentrated on small clusters of spin configurations, those clusters are arranged in clusters, the clusters of clusters are themselved arranged in clusters, and so on. When this clustering stops after k-steps, the system is said to be k-step replica symmetry breaking, otherwise the system exhibits full replica symmetry breaking. In particular, this gives the Gibbs measure an **ultrametric** structure.

Parisi's predictions for the SK model have been shown to hold rigorously. In [8], D. Ruelle developed the Ruelle probability cascades as an explicit description of the thermodynamic limit of the Gibbs measure for the GREM (generalized random energy model), a model constructed with the predicted clustering phenomenon built in, but which was much easier to study. Through the development of invariance properties by E. Bolthausen and A. Sznitman in [9] and stability properties by S. Ghirlanda and F. Guerra in [10] and M. Aizenman and P. Contucci in [11], further connections between the Ruelle probability cascades and the SK model were found. The full RSB free energy formula has been rigorously verified for the SK model by M. Talagrand in his celebrated work [12], and the ultrametric structure of the Gibbs measure was established by D. Panchenko in [13]. A comprehensive summary of the analysis of the SK model can be found in [14]. Many other models of interest have been rigorously shown to exhibit replica symmetry breaking since then (see for example [2, 15]). We ask the following question.

Question 1.2.3. *For a generic interaction, does the Gibbs measure on a random regular graph exhibit replica symmetry breaking at sufficiently low temperatures?*

We do not answer this question, but do prove a suggestive result.

Definition 1.2.4. *For an interaction Φ, we say the replica symmetric free energy*

prediction fails at inverse temperature β if

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}[Z(\beta\Phi)] - \frac{1}{n}\mathbb{E}[\log Z(\beta\Phi)] > 0.$$

We define $\mathcal{RF}$ to be the set of $\Phi \in C(\mathcal{A}^\Gamma, \mathbb{R})$ such that for some $\beta_0 < \infty$, the replica symmetric free energy prediction fails for Φ at inverse temperature β for all $\beta > \beta_0$.

We may now state the main theorem of 2.

Theorem 1.2.5. *The set $\mathcal{RF}$ contains an open, dense subset of $C(\mathcal{A}^\Gamma, \mathbb{R})$.*

Remark 1.2.6. *While 1.2.5 shows the replica symmetric free energy prediction fails generically (in the sense of Baire category), this is not sufficient to claim that a generic system is replica symmetry breaking, nor is it sufficient to show the existence of a phase transition. To prove 1.2.5 we only need rough bounds on the ground state energy for a typical graph, and we have a stronger analysis for a class of atypical graphs that we construct. To show replica symmetry breaking, one would need much finer information about the Gibbs measure on a typical graph. To show the existence of a phase transition in other models, one usually shows the annealed free energy is an analytic function, that the quenched and annealed free energies agree at sufficiently high temperatures, and that they disagree at sufficiently low temperatures. This implies the existence of some temperature where the quenched free energy is not analytic. However in our model, we cannot rule out the possibility that the annealed free energy is non-analytic, and it is also theoretically possible that the quenched and annealed free energies only agree at $\beta = 0$.*

1.2.1 Related works

A common goal in probability theory is to establish universality of some phenomenon. The classical central limit theorem is the prototypical example of this phenomenon.

In the regime of replica symmetry breaking, P. Carmona and Y. Hu in [16] established that the free energy of the SK model with g_{ij} iid according to any distribution with variance 1 finite third moments is the same as the Gaussian case.

In [3] in Remark 1.4 the authors note that it is expected that a generic ferromagnetic model on a non-amenable graph (for example random regular graphs) will enter a spin-glass phase under a zero magnetization constraint. This is a fundamentally different regime from the setting of this dissertation: the zero magnetization constraint is a global constraint, and not expressible in terms of local interactions.

In [15], the authors establish the replica symmetry breaking phenomenon for an extremely wide class of models on random regular hypergraphs, computing the free energy density exactly. While we do not establish replica symmetry breaking in the sense of overlaps, our results apply to interactions beyond nearest neighbor interactions.

1.3 Sofic entropy

The following material, and the material in 3, are adapted from [17].

The topological entropy of a homeomorphism $T : X \to X$ of a compact Hausdorff space X was introduced in [18]. It was generalized to actions of amenable groups via Følner sequences in the 1970s [19] and to certain non-amenable groups via sofic approximations more recently [20]. It plays a major role in the classification and structure theory of topological dynamical systems.

To explain further, suppose Γ is a countable group with identity 1_Γ and $\sigma : \Gamma \to \mathrm{Sym}(V)$ is a map where V is a finite set and $\mathrm{Sym}(V)$ is the group of permutations of V. It is not required that σ is a homomorphism. Let $D \Subset \Gamma$ (the symbol $\Subset$ denotes a finite subset) and $\delta > 0$. Then σ is called

- (D,δ)**-multiplicative** if

$$\#\{v \in V : \ \sigma(gh)v = \sigma(g)\sigma(h)v \ \forall g,h \in D\} > (1-\delta)|V|,$$

- (D,δ)**-trace preserving** if

$$\#\{v \in V : \ \sigma(f)v \neq v \ \forall f \in D \setminus \{1_\Gamma\}\} > (1-\delta)|V|,$$

- (D,δ)**-sofic** if it is both (D,δ)-multiplicative and (D,δ)-trace preserving.

A **sofic approximation** to Γ consists of a sequence $\Sigma = \{\sigma_i\}_{i\in\mathbb{N}}$ of maps $\sigma_i : \Gamma \to \mathrm{Sym}(V_i)$ such that for all finite $D \subset \Gamma$, $\delta > 0$ and all but finitely many i, σ_i is (D,δ)-sofic. A group is **sofic** if it admits a sofic approximation. In this paper we will usually assume $|V_i| = i$.

If Γ acts by homeomorphisms on a compact Hausdorff space X and a sofic approximation Σ to Γ is given then the Σ**-entropy** of the action is a topological conjugacy invariant, denoted by $h_\Sigma(\Gamma \curvearrowright X) \in \{-\infty\} \cup [0,\infty]$. It is also called **sofic entropy** if Σ is understood. It was first defined in [20] where the authors obtain a variational principle connecting it with the previously introduced notion of sofic measure entropy [21]. It is monotone under embeddings and additive under direct products but not monotone under factor maps. See [22] for a survey.

A curious feature of this new entropy is that it may depend on the choice of sofic approximation. This is not always the case; for example, if Γ is amenable then sofic entropy and classical entropy always agree. However, there are examples of actions $\Gamma \curvearrowright X$ by non-amenable groups Γ with sofic approximations Σ_1, Σ_2 satisfying

$$h_{\Sigma_1}(\Gamma \curvearrowright X) = -\infty < h_{\Sigma_2}(\Gamma \curvearrowright X).$$

See [22, Theorem 4.1]. The case $h_{\Sigma_1}(\Gamma \curvearrowright X) = -\infty$ is considered degenerate: it

implies that there are no good models for the action with respect to the given sofic approximation. Until this paper, it was an open problem whether a mixing action could have two different *non-negative* values of sofic entropy. Our main result is:

Theorem 1.3.1 (Theorem 1.1 in [17])**.** *There exists a countable group Γ, a mixing action $\Gamma \curvearrowright X$ by homeomorphisms on a compact metrizable space X and two sofic approximations Σ_1, Σ_2 to Γ such that*

$$0 < h_{\Sigma_1}(\Gamma \curvearrowright X) < h_{\Sigma_2}(\Gamma \curvearrowright X) < \infty.$$

Remark 1.3.2. *The* **range of sofic entropies** *for an action $\Gamma \curvearrowright X$ is the set of all non-negative numbers of the form $h_\Sigma(\Gamma \curvearrowright X)$ as Σ varies over all sofic approximations to Γ. By taking disjoint unions of copies of sofic approximations, it is possible to show the range of sofic entropies is an interval (which may be empty or a singleton). So for the example of Theorem 1.3.1, the range of sofic entropies is uncountable.*

Remark 1.3.3. *It remains an open problem whether there is a measure-preserving action $\Gamma \curvearrowright (X, \mu)$ with two different non-negative sofic entropies. Theorem 1.3.1 does not settle this problem because it is entirely possible that any invariant measure μ on X with $h_{\Sigma_2}(\Gamma \curvearrowright (X, \mu)) > h_{\Sigma_1}(\Gamma \curvearrowright X)$ satisfies $h_{\Sigma_1}(\Gamma \curvearrowright (X, \mu)) = -\infty$.*

1.3.1 Random sofic approximations

We do not know of any explicit sofic approximations to Γ which are amenable to analysis. Instead, we study **random sofic approximations**. For the purposes of this dissertation, these are sequences $\{\mathbb{P}_n\}_n$ of probability measures $\mathbb{P}_n$ on spaces of homomorphisms $\mathrm{Hom}(\Gamma, \mathrm{Sym}(n))$ such that, for any finite $D \Subset \Gamma$ and $\delta > 0$ there is an $\epsilon > 0$ such that

$$\mathbb{P}_n(\sigma \text{ is } (D, \delta)\text{-sofic}) > 1 - n^{-\epsilon n}$$

for all sufficiently large n. Because $n^{-\epsilon n}$ decays super-exponentially, if $\Omega_n \subset \mathrm{Hom}(\Gamma, \mathrm{Sym}(n))$ is any sequence with an exponential lower bound of the form $\mathbb{P}_n(\Omega_n) > e^{-cn}$ (for some constant $c > 0$) then there exists a sofic approximation $\Sigma = \{\sigma_n\}$ with $\sigma_n \in \Omega_n$ for all n.

It is this non-constructive existence result that enables us to use random sofic approximations to prove Theorem 1.3.1.

Chapter 2

The quenched and annealed free energy for generic interactions

The key idea behind the main construction of this chapter is to perturb a given interaction function in order to introduce frustrations. For example, the anti-ferromagnetic Ising model on a graph with an odd cycle is frustrated: there is no way to assign spins on the cycle such that every pair of neighbors has a different spin. We use the fact that w.h.p. a random regular graph is not bipartite to introduce an appropriate frustration that will break the replica symmetric free energy prediction.

To prove 1.2.5, we first define an easier to study subset $\mathcal{G}$, based on the ground state energy $\min_{\chi \in \mathcal{A}^{V_n}} H_\Phi(\chi)$, which we will show is open and dense. That $\mathcal{G}$ is open is very simple to show and is presented in 2.3, only relying on basic Lipschitz properties of the ground state energy. The main contribution is showing that $\mathcal{G}$ is dense in 2.4.

2.1 Non-backtracking matrix of regular graphs

Throughout this chapter we will make use of the non-backtracking matrix of a graph. The adjacency matrix $A_G \in \mathrm{Mat}_{V\times V}(\mathbb{R})$ of a graph $G = (V, E)$ is an intensely studied object. Powers of A count the paths of a given length between vertices,

and the spectrum of this matrix determines expansion properties. In recent years a closely related matrix, the non-backtracking matrix, has received a large amount of attention, for example in Bordenave's simpler proof of Freidman's theorem on the second eigenvalue for the adjacency matrix of random regular graphs [23], or in the spectral redemption results for the stochastic block model [5]. Given an undirected graph $G = (V, E)$ let $\vec{E}$ denote the set of directed edges, two for each original edge in the graph. Each $\vec{e} \in \vec{E}$ then has an origin vertex $o(\vec{e})$ and a target vertex $t(\vec{e})$, as well as a mirror directed edge $\vec{e}^*$. The non-backtracking matrix $B := B_G \in \mathrm{Mat}_{\vec{E}\times\vec{E}}$ is then given by

$$B_{\vec{e},\vec{f}} = \begin{cases} 1 & \text{if } t(\vec{e}) = o(\vec{f}) \text{ and } \vec{f} \neq \vec{e}^* \\ 0 & \text{otherwise} \end{cases}.$$

Powers of this matrix count *non-backtracking* paths between vertices. A path is non-backtracking if it never doubles back across an edge it just traversed.

The adjacency matrix of a graph and its non-backtracking matrix are closely related to one another, particularly for regular graphs.

Proposition 2.1.1 ([?])**.** *Let $G = (V, E)$ be a connected d-regular graph ($d \geq 3$) on n vertices. Let $N = dn$ and let $\{\lambda_i\}_{i=1}^n$ be the eigenvalues of the adjacency matrix, with $\lambda_1 = d$. Then the non-backtracking matrix is unitarily similar to*

$$\Lambda = \mathrm{diag}\left(d-1, \begin{pmatrix} \theta_2 & \alpha_2 \\ 0 & \theta_2' \end{pmatrix}, \cdots, \begin{pmatrix} \theta_n & \alpha_n \\ 0 & \theta_n' \end{pmatrix}, -1, \cdots, -1, 1, \cdots, 1\right)$$

where the eigenvalue -1 occurs with multiplicity $nd/2 - n$, 1 occurs with multiplicity $nd/2 - n + 1$, $|\alpha_i| < 2(d-1)$ for all i, and $\theta_i, \theta_i' \in \mathbb{C}$ are solutions to

$$\theta^2 - \lambda_i \theta + d - 1 = 0.$$

One consequence of this proposition is that for the non-backtracking matrix for a

connected d-regular graph, all eigenvectors for eigenvalues not $d-1$ are orthogonal to $\mathbb{1}$, the vector with every entry 1 and eigenvector with eigenvalue $d-1$.

2.2 Relationship between ground state energy and $\mathcal{RF}$

Definition 2.2.1. *For an interaction $\Phi \in C(\mathcal{A}^\Gamma, \mathbb{R})$, we say the large deviations rate function for the ground state energy is supported on an interval if there exists $\epsilon > 0$ such that*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\min_{\chi\in\mathcal{A}^{V_n}} H_\Phi(\chi) < \mathbb{E}\left[\min_{\chi\in\mathcal{A}^{V_n}} H_\Phi(\chi)\right] - \epsilon n\right] > -\infty. \tag{2.1}$$

We denote the set of such Φ by $\mathcal{G}$.

The connection between $\mathcal{G}$ and $\mathcal{RF}$ is given by the following two results. The arguments are standard, but we include them for completeness.

Claim 2.2.2. *If $\Phi : \mathcal{A}^\Gamma \to \mathbb{R}$ is bounded, then*

$$\lim_{\beta\to\infty} \limsup_{n\to\infty} \left| \frac{1}{\beta n} \mathbb{E}[\log Z(\beta\Phi)] + \frac{1}{n} \mathbb{E}\left[\min_{\chi\in\mathcal{A}^V} H_\Phi(\chi)\right] \right| = 0.$$

Proof. It is clear that

$$e^{-\beta \min H_\Phi} \le \sum_{\chi\in\mathcal{A}^V} e^{-\beta H_\Phi(\chi)} \le |\mathcal{A}|^n e^{-\beta \min H_\Phi}.$$

Taking the logarithm, adding $\beta \min H_\Phi$, taking the expected value, then dividing by βn and taking the appropriate limits yields the desired result since $|\mathcal{A}| < \infty$. □

Lemma 2.2.3. $\mathcal{G} \subseteq \mathcal{RF}$.

Proof. Suppose $\Phi \in \mathcal{G}$. Then Φ satisfies 2.1 with corresponding $\epsilon > 0$. Let us denote

$$C = \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\min H_\Phi < \mathbb{E}[\min H_\Phi] - \epsilon n\right] > -\infty.$$

Then

$$\begin{aligned}
\frac{1}{n} \log \mathbb{E}[Z(\beta\Phi)] &\geq \frac{1}{n} \log \mathbb{E}[e^{-\beta \min H_\Phi}] \\
&\geq \frac{1}{n} \log \left(e^{-\beta\mathbb{E}[\min H_\Phi] + \beta\epsilon n} \mathbb{P}\left[\min H_\Phi < \mathbb{E}[\min H_\Phi] - \epsilon n\right]\right) \\
&= -\frac{\beta}{n}\mathbb{E}[\min H_\Phi] + \beta\epsilon + \frac{1}{n} \log \mathbb{P}[\min H_\Phi < \mathbb{E}[\min H_\Phi] - \epsilon n].
\end{aligned}$$

By 2.2.2 we may choose β_0 and N sufficiently large such that for all $n > N$ and $\beta > \beta_0$

$$\begin{aligned}
&\frac{1}{\beta n}\mathbb{E}[\log Z(\beta\Phi)] < \frac{\epsilon}{2} - \frac{1}{n}\mathbb{E}[\min H_\Phi] \\
&\frac{1}{n} \log \mathbb{P}[\min H_\Phi < \mathbb{E}[\min H_\Phi] - \epsilon n] > C - \frac{\epsilon}{2} \\
&\text{and } \frac{\beta\epsilon}{2} + C > \epsilon.
\end{aligned}$$

Then

$$\begin{aligned}
-\frac{\beta}{n}\mathbb{E}[\min H_\Phi] + \beta\epsilon + \frac{1}{n} \log \mathbb{P}[\min H_\Phi < \mathbb{E}[\min H_\Phi] - \epsilon n] &> \frac{1}{n}\mathbb{E}[\log Z(\beta\Phi)] + \frac{\beta\epsilon}{2} + C - \frac{\epsilon}{2} \\
&> \frac{1}{n}\mathbb{E}[\log Z(\beta\Phi)] + \frac{\epsilon}{2}.
\end{aligned}$$

Finally we see for $n > N$ and $\beta > \beta_0$ that

$$\frac{1}{n} \log \mathbb{E}[Z(\beta\Phi)] - \frac{1}{n}\mathbb{E}[\log Z(\beta\Phi)] > \frac{\epsilon}{2}$$

which implies $\Phi \in \mathcal{RF}$. □

2.3 $\mathcal{G}$ is open

Lemma 2.3.1. *$\mathcal{G}$ is open.*

Proof. Note for any $\chi \in K^{V_n}$ that $\Phi \mapsto H_\Phi(\chi)$ is n-Lipschitz. Thus $\Phi \mapsto \min H_\Phi$ and $\Phi \mapsto \mathbb{E}[\min H_\Phi]$ are n-Lipschitz as well, which implies

$$H_{\Phi+\delta\Psi} - \mathbb{E}[\min H_{\Phi+\delta\Psi}] < H_\Phi - \mathbb{E}[\min H_\Phi] + 2\delta n$$

for any $\Psi \in C(K^\Gamma, \mathbb{R})$ with $\|\Psi\|_\infty \leq 1$ and $\delta > 0$. Suppose Φ satisfies 2.1 with $\epsilon > 0$. Then we have

$$\begin{aligned}&\mathbb{P}\left[\min H_{\Phi+\frac{\epsilon}{3}\Psi} < \mathbb{E}[\min H_{\Phi+\frac{\epsilon}{3}\Psi}] - \frac{\epsilon}{3}n\right] \\ &\geq \mathbb{P}\left[\min H_\Phi < \mathbb{E}[\min H_\Phi] - \epsilon n\right].\end{aligned}$$

Applying the logarithm to both sides and taking the appropriate limits implies $\Phi + \frac{\epsilon}{3}\Psi \in \mathcal{G}$. □

2.4 Construction of interactions in $\mathcal{RF}$

We prove the remaining part of 1.2.5 with the following.

Theorem 2.4.1. *$\mathcal{G}$ is dense.*

We begin by constructing the set of interactions which will act as our perturbations for a general interaction. For an element $s \in \{s_1^{\pm 1}, s_2^{\pm 1}, \cdots, s_d^{\pm 1}\}$, let $[s]$ denote the set of group elements in Γ whose reduced word representation begins with s. Let us define for $\ell \in \mathbb{N}$ and $1 \leq i \leq 2d$

$$B_\ell^{(i)} = \begin{cases} [s_i] \cap B_\ell(1_\Gamma) & 1 \leq i \leq d \\ [s_{i-d}^{-1}] \cap B_\ell(1_\Gamma) & d+1 \leq i \leq 2d \end{cases}.$$

That is, $B_\ell^{(i)}$ for $1 \leq i \leq 2d$ are the branches of the Cayley tree restricted to levels at most ℓ.

Definition 2.4.2. *Given $L \in \mathbb{N}$, $1 \leq i \leq 2d$, and $x \in K^\Gamma$ we define the* ***branch empirical distribution*** *of x to be the probability measure $\mathcal{E}_{i,L}(x) \in \mathrm{Prob}(K)$ given by*

$$\mathcal{E}_{i,L}(x)(k) = \frac{1}{(2d-1)^{L-1}} \left| x^{-1}(k) \cap B_L^{(i)} \setminus B_L^{(i)} \right|.$$

Let $k \in K$, $p, q \in \mathbb{R}$, and $L \in \mathbb{N}$. Then we define $E_{k,L,p,q} \subseteq K^\Gamma$ to be the set of $x \in K^\Gamma$ such that for all $\ell \in \mathbb{N}$ with $L/2 \leq \ell \leq L$ and $1 \leq i \leq 2d$

$$\mathcal{E}_{i,\ell}(x)(k) \in \begin{cases} [p + \frac{2}{3}q, p + \frac{4}{3}q] & \text{if } \ell \equiv 0 \mod 2 \\ [p - \frac{1}{3}q, p + \frac{1}{3}q] & \text{if } \ell \equiv 1 \mod 2 \end{cases}.$$

In words, we require that k has frequency close to $p + q$ on even levels of each branch, and frequency close to p on odd levels. Similarly, define $O_{k,L,p,q}$ with the reversed parity requirement for ℓ instead. We then define $\Psi^{even}_{k,L,p,q} = -\mathbb{1}_{E_{k,L,p,q}}$ and $\Psi^{odd}_{k,L,p,q} = -\mathbb{1}_{O_{k,L,p,q}}$.

Fix $\epsilon > 0$ and an interaction $\Phi \in C(K^\Gamma, \mathbb{R})$. We must find some interaction Ψ such that $\|\Psi\| \leq 1$ and $\Phi + \epsilon\Psi \in \mathcal{G}$. Since K^Γ is compact, without loss of generality we may also assume $||\Phi||_\infty \leq 1$.

For $\delta = \frac{\epsilon}{40} > 0$, let N be sufficiently large such that

$$\frac{1}{N}\mathbb{E}_N[\min H_{\Phi,N}] < \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}_n[\min H_{\Phi,n}] + \delta$$

and such that

$$\mathbb{P}_N[\text{The spectral gap of the non-backtracking matrix of } G_N > .01] > 1 - \frac{\delta}{8}.$$

Such an N exists by Theorem 2 in [23]. Note that $\min H_{\Phi,N} + N$ is a non-negative

random variable since $H_{\Phi,N}(\chi) \geq -N\|\Phi\|_\infty \geq -N$. Thus by Markov's inequality

$$\begin{aligned}\mathbb{P}_N\left[\min H_{\Phi,N} \leq \mathbb{E}[\min H_{\Phi,N}] + \delta N\right] &= 1 - \mathbb{P}_N\left[\min H_{\Phi,N} + N > \mathbb{E}[\min H_{\Phi,N}] + N + \delta N\right] \\ &\geq 1 - \frac{\mathbb{E}_N[\min H_{\Phi,N}] + N}{\mathbb{E}_N[\min H_{\Phi,N}] + N + \delta N} \\ &\geq 1 - \frac{2}{2+\delta} \geq \frac{\delta}{4} > 0.\end{aligned}$$

Thus we may pick some homomorphism $\sigma^* \in \mathrm{Hom}(\Gamma, \mathrm{Sym}(N))$ such that $\frac{1}{N}\min H_{\Phi,N}$ on the graph induced by σ^* is at most $\frac{1}{N}\mathbb{E}_N[\min H_{\Phi,N}] + \delta < \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}_n[\min H_{\Phi,n}] + 2\delta$ and such that the spectral gap of the non-backtracking matrix is larger than .01. Let χ^* denote the minimizer of $H_{\Phi,N}$ (choosing arbitrarily among ties). Pick some $k^* \in K$ such that

$$p^* := \frac{1}{N}(\chi^*)^{-1}(k^*) \leq \frac{1}{2}.$$

Such a spin necessarily exists since $|K| \geq 2$.

We claim

$$\Psi := \Psi^{even}_{k^*,L,p^*,q} + \Psi^{odd}_{k^*,L,p^*,q}$$

for $0 < q < \min\left\{\frac{\epsilon}{40\|\Phi\|_\infty}, 1-p^*\right\}$ and $L > \max\left\{\frac{2\log q/4}{\log(1-1/(100(d-1)))}, 6\right\}$ is our desired interaction. First it is clear that $\Psi \in C(K^\Gamma, \mathbb{R})$ since it is $B_L(1_\Gamma)$-local. It is also clear that $E_{L,p}$ and $O_{L,p}$ are disjoint, so $\|\Psi\|_\infty \leq 1$.

We will now prove 2.4.1 in two parts: first bounding $\mathbb{E}[\min H_{\Phi+\epsilon\Psi}]$, and then constructing a sequence of not too unlikely events where $\min H_{\Phi+\epsilon\Psi} < \mathbb{E}[\min H_{\Phi+\epsilon\Psi}] - \delta n$.

2.4.1 Bounding $\mathbb{E}[\min H_{\Phi+\epsilon\Psi}]$

We will use the following property of the sets $E_{k,L,p,q}$ and $O_{k,L,p,q}$.

Proposition 2.4.3. *If $L \geq 6$, for any $k \in K$, $p, q \in \mathbb{R}_{\geq 0}$, and $1 \leq i \leq d$*

$$s_i E_{k,L,p,q} \cap E_{k,L,p,q} = \emptyset \text{ and } s_i O_{k,L,p,q} \cap O_{k,L,p,q} = \emptyset.$$

Proof. We will prove the proposition for $E_{k,L,p,q}$, the proof for $O_{k,L,p,q}$ is identical. We may assume without loss of generality that $i = 1$. Suppose $s_1 x \in E_{k,L,p,q}$. It is clear that

$$(s_1 x)^{-1}(k) = s_1(x^{-1}(k)) \text{ and } s_1 B_\ell^{(d+1)} = \{1_\Gamma\} \cup \bigcup_{i=2}^{2d} B_{\ell-1}^{(i)}.$$

Thus for any $\ell \geq 3$

$$\begin{aligned}
\left| x^{-1}(k) \cap B_\ell^{(d+1)} \setminus B_{\ell-1}^{(d+1)} \right| &= \left| s_1 \left(x^{-1}(k) \cap B_\ell^{(d+1)} \setminus B_{\ell-1}^{(d+1)} \right) \right| \\
&= \left| (s_1 x)^{-1}(k) \cap \bigcup_{i=2}^{2d} B_{\ell-1}^{(i)} \setminus B_{\ell-2}^{(i)} \right| \\
&= \sum_{i=2}^{2d} \left| (s_1 x)^{-1}(k) \cap B_{\ell-1}^{(i)} \setminus B_{\ell-2}^{(i)} \right|.
\end{aligned}$$

Let $\ell \in 2\mathbb{N}$ be such that $\ell - 1 \geq L/2$ and $\ell \leq L$. Such an ℓ exists since $L \geq 6$. Then

$$\begin{aligned}
\frac{1}{(2d-1)^{\ell-1}} \sum_{i=2}^{2d} \left| (s_1 x)^{-1}(k) \cap B_{\ell-1}^{(i)} \setminus B_{\ell-2}^{(i)} \right| &\leq \frac{1}{(2d-1)} \sum_{i=2}^{2d} p + \frac{1}{3} q \\
&= p + \frac{1}{3} q < p + \frac{2}{3} q.
\end{aligned}$$

Thus

$$\frac{1}{(2d-1)^{\ell-1}} \left| x^{-1}(k) \cap B_\ell^{(d+1)} \setminus B_{\ell-1}^{(d+1)} \right| < p + \frac{2}{3} q$$

and $x \notin E_{k,L,p,q}$. □

We may now analyze the expected value of the minimum of $H_{\Phi + \epsilon \Psi}$.

Claim 2.4.4. *We have*

$$\limsup_{n\to\infty} -\frac{1}{n}\mathbb{E}[\min H_{\Phi+\epsilon\Psi}] \leq \limsup_{n\to\infty} -\frac{1}{n}\mathbb{E}[\min H_{\Phi}] + 2\mathbb{E}[\alpha(G)]\epsilon$$

where $\alpha(G)$ is the maximum density of an independent set in the graph.

Proof. We color a vertex red if $\Psi_{L,p}^{even}(\Pi_v^{\sigma_n}(\chi)) = -1$ and blue if $\Psi_{L,p}^{odd}(\Pi_v^{\sigma_n}(\chi)) = -1$. Clearly vertices cannot be colored both red and blue. Furthermore, two neighboring vertices may not share the same color either by 2.4.3 and the fact that $\Pi^{\sigma_n}_{\sigma_n(s_i)(v)}(\chi) = s_i\Pi_v^{\sigma_n}(\chi)$. Thus the set of red and blue vertices each form independent sets, which implies $H_{\Phi+\epsilon\Psi} \geq H_\Phi - 2\alpha(G)n\epsilon$ from which the claim follows. □

We may bound $\mathbb{E}[\alpha(G)]$ using a simple first moment bound.

Proposition 2.4.5. *For the random permutation model,* $\limsup_{n\to\infty}\mathbb{E}[\alpha(G)] < .45$ *for any* $d \geq 2$.

Proof. Let $Z(\alpha)$ count the number of independent sets of density α. Then $\mathbb{P}[\alpha(G) \geq \alpha] = \mathbb{P}[Z(\alpha) > 0]$, since the existence of an independent set of a given density implies the existence of an independent set for any smaller density. Since $Z(\alpha)$ is an integer valued random variable, $\mathbb{P}[Z(\alpha) > 0] \leq \mathbb{E}[Z(\alpha)]$ by Markov's inequality, and we have

$$\mathbb{E}[Z(\alpha)] = \sum_{\substack{\chi\in\{0,1\}^n \\ \chi^{-1}(1)=\alpha n}} \mathbb{P}[\chi \text{ is an independent set}].$$

By symmetry this sum is equal to

$$\binom{n}{n\alpha}\mathbb{P}[\chi \text{ is an independent set}]$$

where χ is some fixed spin configuration with $\chi^{-1}(1) = \alpha n$. We may construct without repetition all permutations π such that $(\chi(v), \chi(\pi(v)) \neq (1,1)$ for all $v \in V$ by the following process.

1. Choose two subsets V_0 and W_0 of size αn from $\chi^{-1}(0)$.

2. Choose an injective mapping from $\chi^{-1}(1) \to V_0$ and one from $\chi^{-1}(1) \to W_0$.

3. Choose an injective mapping from $\chi^{-1}(0) \setminus V_0 \to \chi^{-1}(0) \setminus W_0$.

This yields

$$\begin{aligned}\mathbb{P}[\chi \text{ is an independent set}] &= \frac{1}{n!^d}\left(\binom{n(1-\alpha)}{n\alpha}^2 n\alpha!^2 n(1-2\alpha)!\right)^d \\ &= \left(\frac{n(1-\alpha)!^2}{n(1-2\alpha)!n!}\right)^d \\ &= \left(\binom{n(1-\alpha)}{n\alpha}\Big/\binom{n}{n\alpha}\right)^d.\end{aligned}$$

Thus by Stirling's approximation

$$\begin{aligned}\mathbb{E}[Z(\alpha)] &= \binom{n(1-\alpha)}{n\alpha}^d \Big/ \binom{n}{n\alpha}^{d-1} \\ &= \exp n\left((1-d)H(\alpha) + d(1-\alpha)H(\alpha/(1-\alpha)) + o(1)\right) \\ &= \exp n\left(H(\alpha) + 2d(1-\alpha)\log(1-\alpha) - d(1-2\alpha)\log(1-2\alpha) + o(1)\right)\end{aligned}$$

where $H(x) = -x\log x - (1-x)\log(1-x)$ is the Shannon entropy. Let f_d denote this exponential growth rate function. Then $f_d(0) = 0$, $f_d'(0) > 0$ and $f_d''(\alpha) < 0$ for all $\alpha \in (0, 1/2]$, so f_d has a unique zero in $(0, 1/2]$. Furthermore, $2(1-\alpha)\log(1-\alpha) - (1-2\alpha)\log(1-2\alpha) \leq 0$ for all $\alpha \in [0, 1/2]$, so it suffices to show that the unique zero of f_2 is at most .45. This follows by simply evaluating $f_2(.45) = -.166\ldots$, which is negative. □

2.4.2 Bounding the probability of $\min H_{\Phi+\epsilon\Psi} < \mathbb{E}[\min H_{\Phi+\epsilon\Psi}] - \delta n$

Recall our choices of N, σ^*, and χ^*. Let $\text{sign} : \Gamma \to \mathbb{Z}/2\mathbb{Z}$ be the homomorphism defined by $\text{sign}(s_1) = \text{sign}(s_2) = \cdots = \text{sign}(s_d) = 1$. We define $\tilde{\sigma} : \Gamma \to \text{Sym}(V_N \times \mathbb{Z}/2\mathbb{Z})$ by

$$\tilde{\sigma}(\gamma)(v, i) = (\sigma^*(\gamma)(v), \text{sign}(\gamma)i).$$

Note that the graph generated by $\tilde{\sigma}$ is the bipartite double cover of the graph generated by σ^*. We also define $\tilde{\chi} \in K^{V_N \times \mathbb{Z}/2\mathbb{Z}}$ to be the lift of χ^*, defined by $\tilde{\chi}(v, i) = \chi^*(v)$ for each $v \in V_N$ and $i \in \mathbb{Z}/2\mathbb{Z}$. We recall some basic facts about the bipartite double cover of a graph G.

- If $\mathcal{B}$ is the non-backtracking matrix of G, then the bipartite double cover has non-backtracking matrix
$$\begin{pmatrix} 0 & \mathcal{B} \\ \mathcal{B} & 0 \end{pmatrix}.$$
- If $\vec{v}$ is an eigenvector of $\mathcal{B}$ with eigenvalue λ, then $\begin{pmatrix} \vec{v} \\ \vec{v} \end{pmatrix}$ and $\begin{pmatrix} \vec{v} \\ -\vec{v} \end{pmatrix}$ are eigenvectors of the bipartite double cover non-backtracking matrix with eigenvalues λ and $-\lambda$ respectively.

Lemma 2.4.6. *If* $(1-1/100(2d-1))^{L/2} < \frac{q}{4}$, *for any* $(v, i) \in V_N \times \mathbb{Z}/2\mathbb{Z}$, $L/2 \leq \ell \leq L$, $1 \leq j \leq 2d$, $\chi \in K^{V_N \times \mathbb{Z}/2\mathbb{Z}}$, *and* $k \in K$

$$\mathcal{E}_{i,\ell}(\chi)(k) \in \begin{cases} \frac{1}{N}\left|\chi^{-1}(k) \cap V_N \times \{i\}\right| + [-q/4, q/4] & \text{if } \ell \equiv 0 \mod 2 \\ \frac{1}{N}\left|\chi^{-1}(k) \cap V_N \times \{i+1\}\right| + [-q/4, q/4] & \text{if } \ell \equiv 1 \mod 2 \end{cases}.$$

Proof. Let $\tilde{\mathcal{B}}$ denote the non-backtracking matrix of the graph generated by $\tilde{\sigma}$. Let $v \in V_N$. Let us assume $i = 0$ and $\ell \equiv 0 \mod 2$, the other cases follow identically.

Fix $1 \le i \le 2d$. Let us denote $\vec{v} = \delta_{((v,0),(\sigma^*(s_i)(v),1))}$. Then clearly

$$\tilde{\mathcal{B}}^{\ell-1}\vec{v} = \sum_{\gamma \in B_\ell^{(j)}} \sum_{m \neq j} \delta_{((\sigma^*(\gamma)(v),0),(\sigma^*(\gamma s_j)(v),1))}.$$

This follows simply from the observation that the non-backtracking matrix on $\tilde{\sigma}$ lifts to the non-backtracking operator on Γ, and that a non-backtracking walk on Γ started from $(1_\Gamma, s_i)$ uniformly distributes over $B_\ell^{(i)}$. Thus

$$\frac{1}{(2d-1)^\ell}\left|\left(\Pi_v^{\tilde{\sigma}}(\chi)\right)^{-1}(k) \cap B_\ell^{(j)} \setminus B_{\ell-1}^{(j)}\right| = \sum_{w,w' \in V_N} \mathbb{1}_{\chi(w)=k^*}\left(\left(\frac{1}{2d-1}\mathcal{B}\right)^{\ell-1}\vec{v}\right)((w,0),(w',1)).$$

Now by assumption on σ^*, $\frac{1}{2d-1}\mathcal{B}$ has top eigenvalues 1 and -1, and every other eigenvalue has magnitude at most $1 - \frac{1}{100(2d-1)}$. We may decompose

$$\vec{v} = \frac{1}{4dN}\begin{pmatrix}\mathbb{1}_{2dN}\\ \mathbb{1}_{2dN}\end{pmatrix} + \frac{1}{4dN}\begin{pmatrix}\mathbb{1}_{2dN}\\ -\mathbb{1}_{2dN}\end{pmatrix} + \vec{v}^{\perp}.$$

Here $\mathbb{1}_n$ is the all 1's vector in $\mathbb{R}^n$ and we identify $\mathbb{R}^{4dN}$ with $\mathbb{R}^{V_N\times\{0\}} \times \mathbb{R}^{V_N\times\{1\}}$. The vector $v^\perp \in \text{span}\left\{\begin{pmatrix}\mathbb{1}_{2dN}\\ \mathbb{1}_{2dN}\end{pmatrix}, \begin{pmatrix}\mathbb{1}_{2dN}\\ -\mathbb{1}_{2dN}\end{pmatrix}\right\}^\perp$ by 2.1.1. It is clear that $\|\vec{v}^\perp\| \le 1$ by a simple computation. Thus

$$\left\|\left(\frac{1}{2d-1}\mathcal{B}\right)^{\ell-1}\vec{v} - \frac{1}{4dN}\begin{pmatrix}\mathbb{1}_{2dN}\\ \mathbb{1}_{2dN}\end{pmatrix} - \frac{1}{4dN}\begin{pmatrix}\mathbb{1}_{2dN}\\ -\mathbb{1}_{2dN}\end{pmatrix}\right\| \le \left(1 - \frac{1}{100(2d-1)}\right)^\ell$$

and since $\frac{1}{4dN}\sum_{w,w' \in V_N} \mathbb{1}_{\chi(w)=k^*}\left(\begin{pmatrix}\mathbb{1}_{2dN}\\ \mathbb{1}_{2dN}\end{pmatrix} - \begin{pmatrix}\mathbb{1}_{2dN}\\ -\mathbb{1}_{2dN}\end{pmatrix}\right) = \frac{1}{N}|\chi^{-1}(k) \cap V_N \times \{0\}|$

we find

$$\left| \frac{1}{(2d-1)^\ell} \left| \left(\Pi_v^{\tilde{\sigma}} \chi\right)^{-1}(k) \cap B_\ell^{(i)} \setminus B_{\ell-1}^{(i)} \right| - \frac{1}{N} \left| \chi^{-1}(k) \cap V_N \times \{0\} \right| \right|$$
$$< \left(1 - \frac{1}{100(2d-1)}\right)^\ell \leq \left(1 - \frac{1}{100(2d-1)}\right)^{L/2} < \frac{q}{4}.$$

□

Let $\tilde{\chi}' \in K^{V_N \times \mathbb{Z}/2\mathbb{Z}}$ be some spin configuration such that

$$\tilde{\chi}'(v,0) = \tilde{\chi}(v,0)$$
$$\left|(\tilde{\chi}')^{-1}(k^*) \cap V_N \times \{1\}\right| = N(p+q)$$
$$d_{\mathrm{Ham}}(\tilde{\chi}, \tilde{\chi}') = q.$$

Such a $\tilde{\chi}'$ clearly exists by reassigning the spins of any subset of size Nq in $V_N \times \{1\} \setminus \tilde{\chi}^{-1}(k^*)$ to k^*. This is possible as by assumption

$$\left|V_N \times \{1\} \setminus \tilde{\chi}^{-1}(k^*)\right| = \left|V_N \setminus (\chi^*)^{-1}(k^*)\right| \geq \frac{1}{2}N.$$

Then by 2.4.6 as well as the Lipschitz property of $H_\Phi(\chi)$ with respect to χ

$$\begin{aligned}
\frac{1}{2N} H_{\Phi+\epsilon\Psi}(\tilde{\chi}')(\tilde{\sigma}) &\leq \frac{1}{2N} H_\Phi(\tilde{\chi})(\tilde{\sigma}) + 2q\|\Phi\| - \epsilon \\
&= \frac{1}{N} H_\Phi(\chi^*)(\sigma^*) - \epsilon + 2q\|\Phi\| \\
&\leq \liminf_{n\to\infty} \frac{1}{n} \mathbb{E}_n[\min H_\Phi] + 2q\|\Phi\| + 2\delta - \epsilon \\
&< \liminf_{n\to\infty} \frac{1}{n} \mathbb{E}_n[\min H_\Phi] + \frac{\epsilon}{20} + \frac{\epsilon}{20} - \epsilon \\
&\leq \liminf_{n\to\infty} \frac{1}{n} \mathbb{E}[\min H_{\Phi+\epsilon\Psi}]
\end{aligned}$$

where this last inequality follows by Claim 2.4.4 and Proposition 2.4.5.

Lemma 2.4.7. *For any fixed* $\sigma_N : \Gamma \to \mathrm{Sym}(N)$

$$\liminf_{k\to\infty} \frac{1}{kN} \mathbb{P}_{kN} \left[\sigma_{kN} \text{ is a } k-\text{lift of } \sigma_N\right] > -\infty.$$

Proof. Let us consider spin configurations $\chi : V_{kN} \to V_N$, that is we color the vertices of the graphs generated by a σ_{kN} by the vertices of of V_N. Note σ_{kN} is a k-lift of σ_N precisely when there exists a spin configuration $\chi : V_{kN} \to V_N$ such that for all $1 \leq i \leq d$ and $v \in V_{kN}$

$$\chi(\sigma_{kN}(s_i)(v)) = \sigma_N(s_i)(\chi(v)).$$

Given a coloring $\chi : V_{kN} \to V_N$, we consider the empirical vertex and edge distributions $\dot{\pi}^\chi \in \mathrm{Prob}(V_N)$ and $\hat{\pi}_i^\chi \in \mathrm{Prob}(V_N \times V_N)$ given by

$$\begin{aligned}
\dot{\pi}^\chi(a) &= \frac{1}{kN} \#\{v \in V_{kN} : \chi(v) = a\} \\
\hat{\pi}_i^\chi(a,b) &= \frac{1}{kN} \#\{v \in V_{kN} : \chi(v) = a, \chi(\sigma_{kN}(s_i)(v)) = b\}.
\end{aligned}$$

Then if we consider the probability distributions $\dot{\pi} \in \mathrm{Prob}(V_N)$ and $\hat{\pi}_i \in \mathrm{Prob}(V_N \times V_N)$ given by

$$\begin{aligned}
\dot{\pi}(v) &= \frac{1}{N} \\
\hat{\pi}_i(v,w) &= \mathbb{1}_{w=\sigma_N(s_i)(v)},
\end{aligned}$$

we have σ_{kN} is a k-lift of σ_N precisely when there is some coloring $\chi : V_{kN} \to V_N$ such that

$$\dot{\pi}^\chi = \dot{\pi} \text{ and } \hat{\pi}_i^\chi = \hat{\pi}_i.$$

Let us denote the number of such colorings by $Z(\dot{\pi}, (\hat{\pi}_i)_{i=1}^d)$. We claim that if $k\dot{\pi}(a) \in$

$\mathbb{Z}$ and $k\hat{\pi}(a,b) \in \mathbb{Z}$ for all $a, b \in \mathcal{A}$ then

$$\liminf_{k\to\infty} \frac{1}{kN} \log \mathbb{E}_{kN}[Z(\dot{\pi}, (\hat{\pi}_i)_{i=1}^d)] > -\infty. \tag{2.2}$$

The total number of spin configurations is $|V_N|^{|V_{kN}|} = N^{kN}$. The lemma then follows from 2.2 since the expectation of $Z(\dot{\pi}, (\hat{\pi}_i)_{i=1}^d)$ is at most N^{kN} times the probability of the event that there is a spin configuration with the given empirical vertex and edge distributions. Now it is a standard computation (see Theorem 3.5 in [22] for example) that

$$\lim_{k\to\infty} \frac{1}{kN} \log \mathbb{E}_{kN}[Z(\dot{\pi}), (\hat{\pi}_i)_{i=1}^d)] = (1-d)H(\dot{\pi}) + \sum_{i=1}^{d} H(\hat{\pi}) > -\infty$$

and 2.2 follows. □

It is a simple observation that if $(\sigma^*)^{\otimes k}$ is a k-lift of σ^*, then $H_{\Phi+\epsilon\Psi}(\chi^k)((\sigma^*)^{\otimes k}) = kH_{\Phi+\epsilon\Psi}(\chi^*)(\sigma^*)$. Thus the bound on $\frac{1}{2N} H_{\Phi+\epsilon\Psi}(\tilde{\chi}')(\tilde{\sigma})$ and 2.4.7 establish 1.2.5.

2.5 Subshifts

We have worked only with permissive shift systems, where every spin may neighbor any other. Many models of interest exhibit hard-core (i.e. topological) constraints however, such as independent sets and colorings. To work with such constraints in our set-up, we consider subshifts over Γ.

Definition 2.5.1. *A subset $X \subset K^\Gamma$ is a **subshift** if it is closed and shift-invariant.*

It is an easy fact that any subshift may be presented by a set of forbidden words. That is, there is a collection of finite subsets $\mathcal{F}$ and spin configurations $y_F \in \mathcal{A}^F$ such that $x \in X$ precisely when $\gamma x|_F \neq y_F$ for each $\gamma \in \Gamma$ and $F \in \mathcal{F}$. Subshifts for which this collection is finite form a particularly important class of subshifts.

Definition 2.5.2. *A closed shift-invariant set $X \subseteq K^\Gamma$ is a **shift of finite type** (abbreviated SFT) if there is some finite set $F \Subset \Gamma$ and some set $\mathcal{W} \subset K^F$ such that $x \in X$ if and only if $(\gamma x)|_F \in \mathcal{W}$.*

When X is an SFT, we will often reduce to the case where it is given by pairwise constraints. This is accomplished by considering the **recoding map**. Given a finite set $F \Subset \Gamma$, we define $\rho^F : K^\Gamma \to (K^F)^\Gamma$ by

$$\rho^F(x)_\gamma = x|_{\gamma^{-1}F}.$$

Then the $\rho^F(X)$ is also an SFT and its constraints are given by a $\{0,1\}$-valued matrix for each generator s_i.

Given an subshift X, we must define an appropriate analogue of $\mathcal{A}^{V_n}$.

Definition 2.5.3. *For a graph $G = (V_n, E_n)$ given by $\sigma_n : \Gamma \to \mathrm{Sym}(V_n)$ and SFT X, the **space of microstates of** X **on** G is*

$$\mathcal{M}_n(X) = \{\chi \in K^{V_n} : \Pi_v^{\sigma_n}(\chi) \in X \text{ for all } v \in V_n\}.$$

In words, we require that each vertex sees a locally allowable spin configuration.

Example 2.5.4. *If $F = \{1_\Gamma, s_1, s_1^{-1}, \cdots, s_d, s_d^{-1}\}$ and $\mathcal{F} = \{\xi \in K^F : \xi(1_\Gamma) = 0 \text{ or } \xi(1_\Gamma) = 1 \text{ and } \xi(\gamma) = 0 \text{ for } \gamma \in F \setminus \{1_\Gamma\}\}$, then the resulting subshift X is the golden mean shift [1], and $\mathcal{M}_n(X)$ consists of all independent sets on the graph G.*

Given an interaction $\Phi \in C(X, \mathbb{R})$ we define the partition function at inverse temperature β by

$$Z_n(\beta\Phi) = \sum_{\chi \in \mathcal{M}_n(X)} e^{-H_\Phi(\chi)}$$

[1] The name golden mean shift comes from the subshift of finite type with the given restrictions over $\mathbb{Z}$. Consider the transformation $x \mapsto \varphi x \mod 1$ with partition $\{[0, 1/\varphi), [1/\varphi, 1)\}$ where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. Then the symbolic shift encoded by this system will consist precisely of binary sequences with no consecutive 1's.

where $H_\Phi(\chi) = \sum_{v \in V_n} \Phi(\Pi_v^{\sigma_n}(\chi))$ as usual.

When adapting the results from the previous sections, two problems arise. First, contrary to the permissive case, $\mathcal{M}_n(X)$ may be empty. For example, if $X \subset \{0,1\}^\Gamma$ is the two point set given by the spin configurations which alternate 0 and 1, then with high probability G is not bipartite and so $\mathcal{M}_n(X)$ is empty. Second, when perturbing the coloring $\tilde{\chi}$ to create $\tilde{\chi}'$, we required being able to substitute a small fraction of spins arbitrarily. This posed no issue in the permissive setting, but we may not generally make this substitution for general subshifts (e.g. the previously given two point shift would require each half of a bipartite graph to be colored with a single color, allowing no substitutions).

2.5.1 Lack of valid microstates

For those n where $\mathcal{M}_n(X)$ is empty with positive probability, $\mathbb{E}[\min_{\chi \in X_n} H_\Phi(\chi)] = \infty$ and $\mathbb{E}[\log Z_n(\beta\Phi)] = -\infty$. Thus if we can show that

$$\liminf_{n\to\infty} \frac{1}{n} \mathbb{P}[\mathcal{M}_n(X) \neq \emptyset] > -\infty, \tag{2.3}$$

the desired inequalities in 1.2.4 and 2.2.1 are immediately satisfied at the given indices, and we may restrict to those n where $\mathcal{M}_n(X)$ is non-empty with probability 1. However, 2.3 does not hold for all choices of SFT. Trivially $X = \emptyset$ is an SFT, and $\mathcal{M}_n(X)$ is then empty for all n with probability 1. However, even non-empty SFTs may have no microstates. An example of an SFT over the free group with no periodic spin configurations is given in Theorem 3.3.5 of [24].

In general, checking if $\mathcal{M}_n(X)$ is non-empty, equivalently if periodic points exist, is hard. Over $\mathbb{Z}$ instead of $\mathbb{F}_d$, this problem is tractable, and for 1-step SFTs corresponds to irreducibilty of the transition matrix. As $\mathbb{F}_d$ is given by a free product of copies of $\mathbb{Z}$, restricting to the following class of SFTs simplifies the problem.

Definition 2.5.5. *A 1-step SFT $X \subseteq \mathcal{A}^\Gamma$ is **isotropic** if there is some $\{0,1\}$-valued matrix $A \in \mathrm{Mat}_{\mathcal{A}\times\mathcal{A}}$ such that $x \in X$ precisely when*

$$A_{x_\gamma, x_{s_i\gamma}} = 1$$

for all $\gamma \in \Gamma$.

Note the golden mean shift and coloring shift are isotropic SFTs.

Lemma 2.5.6. *For any isotropic SFT X, if the transition matrix A is irreducible, there is an integer k such that*

$$\liminf_{n\to\infty} \frac{1}{kn} \mathbb{P}_{kn}[\mathcal{M}_{kn}(X) \neq \emptyset] > -\infty.$$

Proof. From the classical theory of subshifts of finite type over $\mathbb{Z}$, if A is irreducible there is some $p \in \mathcal{A}^{\mathbb{Z}}$ which is an element of the subshift defined by A and is periodic. Let k be the period of p. Define the vertex and edge marginals of p by

$$\dot{\pi}^p(a) = \frac{1}{k} \left|\{0 \le i \le k-1 : p_i = a\}\right|$$
$$\hat{\pi}^p(a,b) = \frac{1}{k} \left|\{0 \le i \le k-1 : p_i = a, p_{i+1} = b\}\right|.$$

Recall from the proof of 2.4.7 the quantity $Z(\dot{\pi}^p, (\hat{\pi}^p)_{i=1}^d)$. By the same reasoning

$$\liminf_{n\to\infty} \frac{1}{kn} \mathbb{E}_{kn}[Z(\dot{\pi}^p, (\hat{\pi}^p)_{i=1}^d)] > -\infty.$$

Any element of $\mathcal{A}^{V_{kn}}$ counted by $Z(\dot{\pi}^p, (\hat{\pi}^p)_{i=1}^d)$ is necessarily a valid spin configuration, and since $\limsup_{n\to\infty} \frac{1}{kn} \log \left|\mathcal{A}^{V_{kn}}\right| = \log \mathcal{A} < \infty$, the claim follows. □

2.5.2 Dobrushin mixing

The second issue of overwriting a fraction of the vertices is remedied with the following property.

Definition 2.5.7. *An subshift X is **Dobrushin-mixing** (or D-mixing) if for any set $A \Subset \Gamma$, there is a set $B \Subset \Gamma$ such that if $x, x' \in X$, there is some $y \in X$ such that $y|_A = x|_A$ and $y|_{\Gamma \setminus B} = x'|_{\Gamma \setminus B}$.*

Example 2.5.8 (Golden mean (independent set) SFT)**.** *Assume without loss of generality that $A = B_L(1_\Gamma)$. Let $B = B_{L+1}(1_\Gamma)$. Then for $x, x' \in X$ define $y \in X$ by $y|_A = x|_A$, $y|_{\Gamma \setminus B} = x'|_{\Gamma \setminus B}$, and $y|_{B \setminus A} = 0^{B \setminus A}$. Then for each $\gamma \in A$ with $y(\gamma) = 1$, if $\gamma' \in \partial\gamma \cap A$ then $y(\gamma') = x(\gamma') = 0$ since $x \in X$, and if $\gamma' \in \partial\gamma \cap (\Gamma \setminus A) \subseteq \partial A$ then $y(\gamma') = 0$. The same holds for $\gamma \in \Gamma \setminus B$. Finally for $\gamma \in B \setminus A$, we have $y(\gamma) = 0$ and the local constraint is automatically satisfied. Thus $y \in X$.*

Example 2.5.9 (k-colorings with $k \geq 2d + 1$)**.** *Take the same A and B as in the previous example. Let $x, x' \in X$ and set $y|_A = x|_A$ and $y|_{\Gamma \setminus B} = x'|_{\Gamma \setminus B}$. We may choose for each $\gamma \in B \setminus A$ some $c \in K$ such that $c \neq x(\partial\gamma \cap A)$ and $c \notin x'(\partial\gamma \cap (\Gamma \setminus B))$. This follows simply from the observation that $|\partial\gamma| = 2d$, and so no matter the coloring of $\partial\gamma$ there is always at least one valid color for γ if $k \geq 2d + 1$. Setting $y(\gamma)$ to the corresponding c then yields a valid coloring.*

Note that for a typical k-coloring of Γ, as L tends to infinity the colors at radius L from the identity do not determine the identity uniquely with high probability as long as $d < \frac{1}{2}k(\log k + \log\log k + 1 + o_k(1))$ (this is freezing transition for random colorings of the tree). Unfortunately, the spin configuration χ^* in our proof may not necessarily exhibit this same statistical behavior, and we require that *any* boundary condition does not determine the root, rather than almost any.

Definition 2.5.10. *Let X be a Dobrushin-mixing subshift, and let B be the set associated to $A = \{1_\Gamma\}$. The **radius of interaction** of X is the minimum R such that*

$B \subseteq B_R(1_\Gamma)$.

2.6 Construction of interactions in $\mathcal{RF}$ for subshifts

The goal of this section to prove an analogue of 1.2.5 for subshifts. We first define an analogue of $\mathcal{RF}$.

Definition 2.6.1. *For a subshift X we define $\mathcal{RF}(X)$ to be the set of interactions $\Phi \in C(X, R)$ such that*

$$\liminf_{k\to\infty} \frac{1}{n_k} \log \mathbb{E}_{n_k}[Z_{n_k}(\beta\Phi)] - \frac{1}{n_k}\mathbb{E}_{n_k}[\log Z_{n_k}(\beta\Phi)] > 0$$

where $(n_k)_{k=1}^{\infty}$ is some increasing sequence of natural numbers with positive density.

Due to the possible emptiness of $\mathcal{M}_n(X)$ with probability 1, we cannot require this lim inf holds when taken over all n. As we will require X is a subshift of finite type for the construction as is, we are able to establish the lim inf on a set of positive density.

Theorem 2.6.2. *There is some function $d_\star : \mathbb{N} \to \mathbb{N}$ such that the following holds. Let X is a Dobrushin-mixing subshift of finite type with at least one periodic point on Γ with interaction radius R such that $d \geq d_\star(R)$. Then $\mathcal{RF}(X)$ contains a dense open subset of $C(X, \mathbb{R})$.*

We will first consider a homomorphism $\tau : \Gamma \to \mathrm{Sym}(V_M)$ for $M \in \mathbb{N}$ which will function as the analogue of the sign homomorphism. Specifically, rather than lifting $\sigma : \Gamma \to \mathrm{Sym}(V_N)$ to its bipartite cover, we will lift to $\sigma \times \tau$. We may choose M sufficiently large such that there exists $\tau : \Gamma \to \mathrm{Sym}(V_M)$ such that the non-backtracking matrix of τ is larger than .01, and such that there is some vertex

$w^* \in V_M$ such that $|\sigma(B_R(1_\Gamma))(w^*)| = |B_R(1_\Gamma)|$ where R is the interaction radius of X. That is, that the depth R neighborhood of w^* in the graph is a tree. Both of these events occur with high probability, so for M sufficiently large such a τ exists. Let us denote this homomorphism by τ^*.

For a homomorphism $\sigma : \Gamma \to \mathrm{Sym}(V_N)$, the adjacency matrix of the graph generated by $\sigma \times \tau^*$ is given by $\sum_{i=1}^{d} S_i(\sigma) \otimes S_i(\tau^*) + S_i(\sigma)^* \otimes S_i(\tau^*)^*$ where $S_i(\tau^*) \in \mathrm{Mat}_{M\times M}$ is the permutation matrix associated to $\tau^*(s_i)$. Thus by Theorem 2 in [25] and Corollary 2 in [26], each eigenvalue of the adjacency matrix will, with high probability in σ, have magnitude strictly less than $2d - .01$, except perhaps for the eigenvalues coming from τ^*. By construction, these will also be less than $2d - .01$, and so the adjacency matrix for $\sigma \times \tau^*$ will be greater than .01 with high probability.

Fix $\Phi \in C(X, \mathbb{R})$ and $\epsilon > 0$. Since X is compact, we may assume without loss of generality that $\|\Phi\| \leq 1$. Let us fix $\delta > 0$ sufficiently small in terms of ϵ. Choose N sufficiently large such that $\mathbb{E}[\min_{\chi\in\mathcal{M}_N(X)} H_\Phi(\chi)] > \limsup_{n\to\infty} \mathbb{E}[\min_{\chi\in\mathcal{M}_n(X)} H_\Phi(\chi)] - \delta$ where the lim sup is taken over all n for which $\mathcal{M}_n(X)$ is non-empty with probability 1. We may also take N be sufficiently large such that with probability at least $1-\delta/8$, the adjacency matrix of the graph generated by $\sigma \times \tau^*$ has spectral gap at least .01, and such that this graph is connected. Arguing as we did for the permissive case, we may then select some σ^*, χ^*, and k^* in the exact same way.

As the action of Γ on V_M given by τ^* is transitive, which follows from connectedness of $G(\tau^*)$, we may identify V_M with cosets of $H = \bigcap_{w\in V_M} \mathrm{Stab}(w)$, and the graph $G(\tau^*)$ with the Schreier coset graph of H. Note that, by our choice of τ, H does not intersect $B_R(1_\Gamma)$.

Now let $\tilde{\chi} : V_N \times V_M \to K$ be the spin configuration defined by $\tilde{\chi}(v, w) = \chi^*(v)$ for all $v \in V_N$ and $w \in V_M$. Note that $\tilde{\chi} \in \mathcal{M}_{NM}(X)$ since $\chi \in \mathcal{M}_N(X)$. We construct a spin configuration $\tilde{\chi}' \in \mathcal{M}_{NM}(X)$ by the following procedure. First, pick a $v \in (\chi^*)^{-1}(k^*)$ arbitrarily consider $x = \Pi^{\sigma^*\times\tau^*}_{(v,w^*)}(\tilde{\chi})$. By the Dobrushin-mixing

condition, we may find some valid spin configuration $x' \in X$ such that $x'|_{\Gamma \setminus B_R(1_\Gamma)} = x|_{\Gamma \setminus B_R(1_\Gamma)}$ and $x'(1_\Gamma) \neq x(1_\Gamma)$. Furthermore, since X is an SFT, by shift-invariance and the locality of the validity check, we may make these same substitutions at each $\gamma \in H$. As X is compact, we may take the limit of the spin configurations created by these substitutions, and get an element of X which maps down to a consistent spin configuration on $G(\sigma \times \tau)$. Let us call this configuration $\tilde{\chi}'$.

The key property of $\tilde{\chi}'$ that we will use is that $\tilde{\chi}'|_{V_N \times (V_M \setminus B_R(w^*)} = \tilde{\chi}|_{V_N \times (V_M \setminus B_R(w^*)}$, and that $|(\tilde{\chi}')^{-1}(k^*) \cap (V_N \times \{w^*\})| \neq |\tilde{\chi}^{-1}(k^*) \cap (V_N \times \{w^*\}|$. That is, on the copies of V_N which are sufficiently far from $V_N \times w^*$, $\tilde{\chi}'$ and $\tilde{\chi}$ are identical, while they have a different frequency of the spin k^* on $V_N \times \{w^*\}$.

For $\gamma H \in \Gamma/H$, $L \in \mathbb{N}$, and $\mathcal{O} \subseteq \text{Prob}(\mathcal{A})$ let $S_{\gamma H}(\mathcal{O}, L)$ denote the set of all $x \in X$ such that for all $k \in \mathcal{A}$ and $L/2 \leq \ell \leq L$ with $B_\ell^{(i)} \cap \gamma H \neq \emptyset$

$$\frac{1}{\left|B_\ell^{(i)} \cap \gamma H\right|} \left|x^{-1}(\cdot) \cap B_\ell^{(i)} \cap \gamma H\right| \in \mathcal{O}.$$

We define

$$\Psi_{base} = \prod_{\gamma H \in \Gamma/H} \mathbb{1}_{S_{\gamma H}(\mathcal{O}_{\gamma H}, L)}$$

where

$$\mathcal{O}_{\gamma H} = \left\{ m \in \text{Prob}(\mathcal{A}) : \left| m(k^*) - \frac{1}{N} \left|(\tilde{\chi}')^{-1}(k^*) \cap V_N \times \{\gamma H\}\right| \right| < \frac{1}{2N} \right\}.$$

In words, we require that the empirical branch distributions restricted to the cosets of H closely match those of $\tilde{\chi}'$. We define our perturbation Ψ to be

$$\Psi = \sum_{\gamma H \in \Gamma/H} \Psi_{base} \circ \gamma.$$

We claim that $\|\Psi\| \leq 1$ and that $\Psi_{base}(x) \neq \Psi_{base}(s_i x)$ for any $1 \leq i \leq d$.

That $\|\Psi\| \leq 1$ follows from the observation that for any $\gamma H \in \Gamma/H$, there is some pair $w, w' \in V_M$ such that $\tau(\gamma)(w) = w'$ and $|\tilde{\chi}'(k^*) \cap (V_N \times \{w\})| \neq |\tilde{\chi}'(k^*) \cap (V_N \times \{w'\})|$. We may see this inductively by starting with w^* and $\tau(\gamma)(w^*)$. If the level sets of V_N have different frequencies of k^* under $\tilde{\chi}'$ we are done, otherwise may consider $\tau(\gamma)(w^*)$ and $\tau(\gamma^2)(w^*)$ and so on. This process will eventually reach $V_M \setminus B_R(w^*)$ as H does not intersect $B_R(1_\Gamma)$, at which point we know by construction of $\tilde{\chi}'$ that the frequencies will differ.

That $\Psi_{base}(x) \neq \Psi_{base}(s_i x)$ follows identically to 2.4.3.

Thus, the expected value of the ground state energy of $\Phi + \epsilon\Psi$ is bounded by the expected ground state energy of Φ plus ϵ times densest coloring of vertices by $\Gamma/H \cup \{*\}$ where $*$ is a joker color which may neighbor itself. By a first moment argument, this density is strictly less than 1 if $|\Gamma/H| < d$ (in fact we only require that $|G/H|$ is smaller than the typical colorability threshold for the random permutation model, but this requires a stronger argument than a first moment bound). This requirement is what leads to the function $d_\star$.

On the other hand, picking L sufficiently large and using the spectral gap of the back-tracking matrix of $G(\sigma \times \tau)$ (which follows from the spectral gap of the adjacency matrix of $G(\sigma \times \tau)$ and 2.1.1), we may prove an analogue of 2.4.6, and we see for each $(v, w) \in V_N \times V_M$ that $\Psi(\Pi^{\sigma^* \times \tau^*}_{(v,w)}(\tilde{\chi}')) = 1$ precisely because Ψ is defined in terms of $\tilde{\chi}'$.

The same argument can then be used to bound $\frac{1}{NM} H_{\Phi+\epsilon\Psi}(\tilde{\chi}')$, and taking random lifts of $\sigma^* \times \tau^*$ then yields the required event with probability decreasing at most exponentially.

Chapter 3

Dependence of topological sofic entropy on the choice of sofic approximation

3.0.1 Proper colorings of random hyper-graphs from a statistical physics viewpoint

The idea for our main construction comes from studies of proper colorings of random hyper-graphs. Although these studies have very different motivations than those that inspired this paper, the examples that they provide are roughly the same as the examples used to prove Theorem 1.3.1. The relevant literature and an outline is presented next.

A **hyper-graph** is a pair $G = (V, E)$ where E is a collection of subsets of V. Elements of E are called **hyper-edges** but we will call them **edges** for brevity's sake. G is k**-uniform** if every edge $e \in E$ has cardinality k.

A **2-coloring** of G is a map $\chi : V \to \{0, 1\}$. An edge $e \in E$ is **monochromatic** for χ if $|\chi(e)| = 1$. A coloring is **proper** if it has no monochromatic edges.

Let $H_k(n,m)$ denote a hyper-graph chosen uniformly among all $\binom{\binom{n}{k}}{m}$ k-uniform hyper-graphs with n vertices and m edges. We will consider the number of proper 2-colorings of $H_k(n,m)$ when k is large but fixed, and the ratio of edges to vertices $r := m/n$ is bounded above and below by constants.

This random hyper-graph model was studied in [27, 28, 29]. These works are motivated by the *satisfiability conjecture*. To explain, the **lower satisfiability threshold** $r^-_{\text{sat}} = r^-_{\text{sat}}(k)$ is the supremum over all r such that

$$\lim_{n\to\infty} \text{Prob}[H_k(n, \lceil rn \rceil) \text{ is properly 2-colorable}] = 1.$$

The **upper satisfiability threshold** $r^+_{\text{sat}} = r^+_{\text{sat}}(k)$ is the infimum over all r such that

$$\lim_{n\to\infty} \text{Prob}[H_k(n, \lceil rn \rceil) \text{ is properly 2-colorable}] = 0.$$

The satisfiability conjecture posits that $r^-_{\text{sat}} = r^+_{\text{sat}}$. It is still open.

Bounds on these thresholds were first obtained in [27] as follows. Let $Z(G)$ be the number of proper 2-colorings of a hyper-graph G. A first moment computation shows that

$$f_k(r) = \lim_{n\to\infty} n^{-1} \log \mathbb{E}[Z(H_k(n, \lceil rn \rceil))]$$

where $f_k(r) := \log(2) + r\log(1 - 2^{1-k})$. Let $r_{\text{first}} = r_{\text{first}}(k)$ be such that $f_k(r_{\text{first}}) = 0$. If $r > r_{\text{first}}$ then $f_k(r) < 0$. Therefore $r^+_{\text{sat}} \le r_{\text{first}}$.

Let r_{second} be the supremum over numbers $r \ge 0$ such that the second moment $\mathbb{E}[Z(H_k(n, \lceil rn \rceil))^2]$ is equal to $\mathbb{E}[Z(H_k(n, \lceil rn \rceil))]^2$ up to sub-exponential factors. The Paley-Zygmund inequality gives the bound $r_{\text{second}} \le r^-_{\text{sat}}$.

In [27], it is shown that

$$\begin{aligned} r_{\text{first}} &= \frac{\log(2)}{2}2^k - \frac{\log(2)}{2} + O(2^{-k}), \\ r_{\text{second}} &= \frac{\log(2)}{2}2^k - \frac{\log(2)+1}{2} + O(2^{-k}). \end{aligned}$$

So there is a constant-sized gap between the two thresholds.

A more detailed view of the second moment is illuminating. But before explaining, we need some terminology. Let $[n]$ be the set of natural numbers $\{1, 2, .., n\}$. A coloring χ of $[n]$ is **equitable** if $|\chi^{-1}(0)| = |\chi^{-1}(1)|$. We will assume from now on that n is even so that equitable colorings of $[n]$ exist. Let $Z_e(G)$ be the number of equitable proper colorings of a hyper-graph G. A computation shows that $\mathbb{E}[Z(H_k(n, \lceil rn \rceil))]$ equals $\mathbb{E}[Z_e(H_k(n, \lceil rn \rceil))]$ up to sub-exponential factors. This enables us to work with equitable proper colorings in place of all proper colorings. This reduces the computations because there is only one equitable coloring up to the action of the symmetric group $\text{Sym}(n)$.

A computation shows that the second moment factorizes as

$$\mathbb{E}[Z_e(H_k(n, m))^2] = \mathbb{E}[Z_e(H_k(n, m))]\mathbb{E}[Z_e(H_k(n, m))|\chi \text{ is proper}]$$

where $\chi : [n] \to \{0, 1\}$ is any equitable 2-coloring. Let $H_k^\chi(n, m)$ be the random hyper-graph chosen by conditioning $H_k(n, m)$ on the event that χ is a proper 2-coloring. This is called the **planted model** and χ is the **planted coloring**. So computing the second moment of $Z_e(H_k(n, m))$ reduces to computing the first moment of $Z_e(H_k^\chi(n, m))$.

The **normalized Hamming distance** between colorings $\chi, \chi' : [n] \to \{0, 1\}$ is

$$d_n(\chi, \chi') = n^{-1}\#\{v \in [n] : \ \chi(v) \neq \chi'(v)\}.$$

Let $Z^\chi(\delta)$ be the number of equitable proper colorings χ' with $d_n(\chi, \chi') = \delta$. Then

$$Z_e(H_k^\chi(n, m)) = \sum_\delta Z^\chi(\delta).$$

In [27], it is shown that $\mathbb{E}[Z^\chi(\delta)|\chi \text{ is proper}]$ is equal to $\exp(n\psi(\delta))$ (up to sub-exponential factors) where ψ is an explicit function.

Note that $\psi(\delta) = \psi(1-\delta)$ (since if χ' is a proper equitable coloring then so is $1-\chi'$ and $d_n(\chi, 1-\chi') = 1 - d_n(\chi, \chi')$). A computation shows $\psi(1/2) = f_k(r)$. If $r < r_{\text{second}}$ then $\psi(\delta)$ is uniquely maximized at $\delta = 1/2$. However, if $r > r_{\text{second}}$ then the maximum of ψ is attained in the interval $\delta \in [0, 2^{-k/2}]$. In fact, $\psi(\delta)$ is negative for $\delta \in [2^{-k/2}, 1/2 - 2^{-k/2}]$. So with high probability, there are no proper equitable colorings χ' with $d_n(\chi, \chi') \in [2^{-k/2}, 1/2 - 2^{-k/2}]$. This motivates defining the **local cluster**, denoted $\mathcal{C}(\chi)$, to be the set of all proper equitable 2-colorings χ' with $d_n(\chi, \chi') \leq 2^{-k/2}$.

The papers [28, 29] obtain a stronger lower bound on the lower satisfiability threshold using an argument they call the *enhanced second moment method*. To explain, we need some terminology. We say a proper equitable coloring χ is **good** if the size of the local cluster $|\mathcal{C}(\chi)|$ is bounded by $\mathbb{E}[Z_e(H_k(n, m))]$. One of the main results of [28, 29] is that $\Pr[\chi \text{ is good}|\chi \text{ is proper}]$ tends to 1 as $n \to \infty$ with $m = rn + O(1)$ and $r < r_{\text{second}} + \frac{1-\log(2)}{2} + o_k(1)$. An application of the Paley-Zygmund inequality to the number of good colorings yields the improved lower bound

$$r_{\text{second}} + \frac{1 - \log(2)}{2} + o_k(1) \leq r_{\text{sat}}^-.$$

The argument showing $\Pr[\chi \text{ is good}|\chi \text{ is proper}] \to 1$ is combinatorial. It is shown that (with high probability) there is a set $R \subset [n]$ with cardinality $|R| \approx (1 - 2^{-k})n$ which is rigid in the following sense: if $\chi' : [n] \to \{0, 1\}$ is any proper equitable 2-coloring then either: the restriction of χ' to R is the same as the restriction of χ to R

or $d_n(\chi', \chi)$ is at least cn/k^t for some constants $c, t > 0$. This rigid set is constructed explicitly in terms of local combinatorial data of the coloring χ on $H_k^\chi(n, m)$.

In summary, these papers study two random models $H_k(n, m)$ and $H_k^\chi(n, m)$. When $r = m/n$ is in the interval $(r_{\text{second}}, r_{\text{second}} + \frac{1-\log(2)}{2})$, the typical number of proper colorings of $H_k(n, m)$ grows exponentially in n but is smaller (by an exponential factor) than the expected number of proper colorings of $H_k^\chi(n, m)$. It is these facts that we will generalize, by replacing $H_k(n, m), H^\chi(n, m)$ with random sofic approximations to a group Γ so that the exponential growth rate of the number of proper colorings roughly corresponds with sofic entropy.

Although the models that we study in this paper are similar to the models in [27, 28, 29], they are different enough that we develop all results from scratch. Moreover, although the strategies we employ are roughly same, the proof details differ substantially. The reader need not be familiar with these papers to read this paper.

3.0.2 The action

In the rest of this introduction, we introduce the action $\Gamma \curvearrowright X$ in Theorem 1.3.1 and outline the first steps of its proof. So fix positive integers k, d. Let

$$\Gamma = \langle s_1, \ldots, s_d : \ s_1^k = s_2^k = \cdots = s_d^k = 1 \rangle$$

be the free product of d copies of $\mathbb{Z}/k\mathbb{Z}$.

The **Cayley hyper-tree** of Γ, denoted $G = (V, E)$, has vertex set $V = \Gamma$. The edges are the left-cosets of the generator subgroups. That is, each edge $e \in E$ has the form $e = \{gs_i^j : \ 0 \leq j \leq k-1\}$ for some $g \in \Gamma$ and $1 \leq i \leq d$.

Remark 3.0.1. *It can be shown by considering each element of* Γ *as a reduced word in the generators* $s_1, ..., s_d$ *that* G *is a hyper-tree in the sense that there exists a unique "hyper-path" between any two vertices. More precisely, for any* $v, w \in V$*, there exists*

a unique sequence of edges $e_1, .., e_l$ *such that* $v \in e_1$, $w \in e_l$, $|e_i \cap e_{i+1}| = 1$, $e_i \neq e_j$ *for any* $i \neq j$, *and* $v \notin e_2$, $w \notin e_{l-1}$. *More intuitively, there are no "hyper-loops" in* G.

The group Γ acts on $\{0,1\}^\Gamma$ by $(gx)_f = x_{g^{-1}f}$ for $g, f \in \Gamma, x \in \{0,1\}^\Gamma$. Let $X \subset \{0,1\}^\Gamma$ be the subset of proper 2-colorings. It is a closed Γ-invariant subspace. Furthermore, $\Gamma \curvearrowright X$ is topologically mixing:

Claim 3.0.2. *For any nonempty open sets* A, B *in* X, *there exists* N *such that for any* $g \in \Gamma$ *with* $|g| > N$, $gA \cap B \neq \emptyset$. *Here* $|g|$ *denotes the shortest word length of representations of* g *by generators* $s_1, ..., s_d$.

Proof. It suffices to show the claim for A,B being cylinder sets. We make a further simplification by assuming each A, B is a cylinder set on a union of hyper-edges, and a yet further simplification that each A, B is a cylinder set on a connected union of hyper-edges. Informally, by shifting the "coordinates" on which A depends so that they are far enough separated from the coordinates on which B depends, we can always fill in the rest of the graph to get a proper coloring.

More precisely, suppose $A = \{x \in X : x \restriction F_A = \chi_A\}$, where $\mathcal{F}_A \subset E$ such that for any $e_1, e_2 \in \mathcal{F}_A$ there is a finite sequence of $f_1, \cdots, f_\ell \in \mathcal{F}_A$ such that $e_1 \cap f_1 \neq \emptyset$, $f_i \cap f_{i+1} \neq \emptyset$, and $f_\ell \cap e_2 \neq \emptyset$, $F_A = \cup_{e \in \mathcal{F}_A} e$, and $\chi_A : F_A \to \{0,1\}$ and similarly $B = \{x \in X : x \restriction F_B = \chi_B\}$. χ_A and χ_B must be bichromatic on each edge in their respective domains since A and B are nonempty.

Let $N = \max\{|h| : h \in F_A\} + \max\{|h| : h \in F_B\} + k$. Then it can be shown for any g with $|g| > N$, that $g^{-1}F_A \cap F_B = \emptyset$. It follows from our earlier remark that there exists a unique hyper-path connecting $g^{-1}F_A$ to F_B (otherwise there would be a hyperloop in G). Thus for example one can recursively fill in a coloring on the rest of Γ by levels of hyperedges - first the hyperedges adjacent to $g^{-1}F_A$ and F_B, then the next layer of adjacent hyperedges, and so on. At each step, most hyperedges only

have one vertex whose color is determined, so it is always possible to color another vertex of an edge to make it bichromatic. Only along the hyper-path connecting $g^{-1}F_A$ to F_B at some step there will be a hyperedge with two vertices whose colors are already determined, but since k is large there is still another vertex to color to make the edge bichromatic.

□

We will show that for certain values of k, d, the action $\Gamma \curvearrowright X$ satisfies the conclusion of Theorem 1.3.1.

3.0.3 Sofic entropy of the shift action on proper colorings

Given a homomorphism $\sigma : \Gamma \to \mathrm{Sym}(V)$, let $G_\sigma = (V, E_\sigma)$ be the hyper-graph with vertices V and edges equal to the orbits of the generator subgroups. That is, a subset $e \subset V$ is an edge if and only if $e = \{\sigma(s_i^j)v\}_{j=0}^{k-1}$ for some $1 \leq i \leq d$ and $v \in V$.

Recall that a hyper-graph is **k-uniform** if every edge has cardinality k. We will say that a homomorphism $\sigma : \Gamma \to \mathrm{Sym}(V)$ is **uniform** if G_σ is k-uniform. Equivalently, this occurs if for all $1 \leq i \leq d$, $\sigma(s_i)$ decomposes into a disjoint union of k-cycles.

A 2-coloring $\chi : V \to \{0, 1\}$ of a hyper-graph G is **ϵ-proper** if the number of monochromatic edges is $\leq \epsilon|V|$. Using the formulation of sofic entropy in [22] (which was inspired by [30]), we show in §3.1 that if $\Sigma = \{\sigma_n\}_{n\geq 1}$ is a sofic approximation to Γ by uniform homomorphisms then the Σ-entropy of $\Gamma \curvearrowright X$ is:

$$h_\Sigma(\Gamma \curvearrowright X) := \inf_{\epsilon>0} \limsup_{i\to\infty} |V_i|^{-1} \log \#\{\epsilon\text{-proper 2-colorings of } G_{\sigma_i}\}.$$

3.0.4 Random hyper-graph models

Definition 3.0.3. *Let* $\mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$ *denote the set of all uniform homomorphisms from* Γ *to* $\mathrm{Sym}(n)$. *Let* $\mathbb{P}_n^u$ *be the uniform probability measure on* $\mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$

and let $\mathbb{E}_n^u$ *be its expectation operator. The measure* $\mathbb{P}_n^u$ *is called the* **uniform model**. *We will always assume* $n \in k\mathbb{Z}$ *so that* $\mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$ *is non-empty. In §3.2 we show that* $\{\mathbb{P}_n^u\}_{n\geq 1}$ *is a random sofic approximation. We will use the uniform model to obtain the sofic approximation* Σ_1 *which appears in Theorem 1.3.1.*

Recall that if V is a finite set, then a 2-coloring $\chi : V \to \{0, 1\}$ is **equitable** if $|\chi^{-1}(0)| = |\chi^{-1}(1)|$. We assume from now on that n is even so that equitable colorings of $[n]$ exist.

Definition 3.0.4. *Fix an equitable coloring* $\chi : [n] \to \{0, 1\}$. *Let* $\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ *be the set of all uniform homomorphisms* $\sigma : \Gamma \to \mathrm{Sym}(n)$ *such that* χ *is proper as a coloring on* G_σ. *Let* $\mathbb{P}_n^\chi$ *be the uniform probability measure on* $\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ *and let* $\mathbb{E}_n^\chi$ *be its expectation operator. The measure* $\mathbb{P}_n^\chi$ *is called the* **planted model** *and* χ *is the* **planted coloring**. *When* χ *is understood, we will write* $\mathbb{P}_n^p$ *and* $\mathbb{E}_n^p$ *instead of* $\mathbb{P}_n^\chi$ *and* $\mathbb{E}_n^\chi$. *In §3.2 we show that* $\{\mathbb{P}_n^p\}_{n\geq 1}$ *is a random sofic approximation. We will use the planted model to obtain the sofic approximation* Σ_2 *which appears in Theorem 1.3.1.*

Remark 3.0.5. *If* χ *and* χ' *are both equitable* 2*-colorings then there are natural bijections from* $\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ *to* $\mathrm{Hom}_{\chi'}(\Gamma, \mathrm{Sym}(n))$ *as follows. Given a permutation* $\pi \in \mathrm{Sym}(n)$ *and* $\sigma : \Gamma \to \mathrm{Sym}(n)$, *define* $\sigma^\pi : \Gamma \to \mathrm{Sym}(n)$ *by* $\sigma^\pi(g) = \pi\sigma(g)\pi^{-1}$. *Because* χ *and* χ' *are equitable, there exists* $\pi \in \mathrm{Sym}(n)$ *such that* $\chi = \chi' \circ \pi$. *The map* $\sigma \mapsto \sigma^\pi$ *defines a bijection from* $\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ *to* $\mathrm{Hom}_{\chi'}(\Gamma, \mathrm{Sym}(n))$. *Moreover* π *defines an hyper-graph-isomorphism from* G_σ *to* G_{σ^π}. *Therefore, any random variable on* $\mathrm{Hom}(\Gamma, \mathrm{Sym}(n))$ *that depends only on the hyper-graph* G_σ *up to hyper-graph-isomorphism has the same distribution under* $\mathbb{P}_n^\chi$ *as under* $\mathbb{P}_n^{\chi'}$. *This justifies calling* $\mathbb{P}_n^\chi$ the *planted model.*

3.0.5 The strategy and a key lemma

The idea behind the proof of Theorem 1.3.1 is to show that for some choices of (k, d), the uniform model admits an exponential number of proper 2-colorings, but it has exponentially fewer proper 2-colorings than the expected number of proper colorings of the planted model (with probability that decays at most sub-exponentially in n).

To make this strategy more precise, we introduce the following notation. Let $Z(\epsilon; \sigma)$ denote the number of ϵ-proper 2-colorings of G_σ. A coloring is **σ-proper** if it is $(0, \sigma)$-proper. Let $Z(\sigma) = Z(0; \sigma)$ be the number of σ-proper 2-colorings.

In §3.2, the proof of Theorem 1.3.1 is reduced to the Key Lemma:

Lemma 3.0.6 (Key Lemma)**.** *Let $f(d, k) := \log(2) + \frac{d}{k}\log(1 - 2^{1-k})$. Also let $r = d/k$. Then*

$$f(d,k) \;=\; \lim_{n\to\infty} n^{-1}\log \mathbb{E}_n^u[Z(\sigma)] = \inf_{\epsilon>0}\limsup_{n\to\infty} n^{-1}\log \mathbb{E}_n^u[Z(\epsilon;\sigma)]. \tag{3.1}$$

Moreover, for any

$$0 < \eta_0 < \eta_1 < (1 - \log 2)/2$$

there exists k_0 (depending on η_0, η_1) such that for all $k \geq k_0$ if

$$r = d/k = \frac{\log(2)}{2}\cdot 2^k - (1 + \log(2))/2 + \eta$$

for some $\eta \in [\eta_0, \eta_1]$ then

$$f(d,k) \;<\; \liminf_{n\to\infty} n^{-1}\log \mathbb{E}_n^p[Z(\sigma)]. \tag{3.2}$$

Also,

$$0 \;=\; \inf_{\epsilon>0}\liminf_{n\to\infty} n^{-1}\log\left(\mathbb{P}_n^u\left(\left|n^{-1}\log Z(\sigma) - f(d,k)\right| < \epsilon\right)\right). \tag{3.3}$$

In all cases above, the limits are over $n \in 2\mathbb{Z} \cap k\mathbb{Z}$.

Equations (3.1) and (3.2) are proven in §3.3 and §3.4 using first and second moment arguments respectively. This part of the paper is similar to the arguments used in [27].

Given $\sigma : \Gamma \to \mathrm{Sym}(n)$ and $\chi : [n] \to \{0,1\}$, let $\mathcal{C}_\sigma(\chi)$ be the set of all proper equitable colorings $\chi' : [n] \to \{0,1\}$ with $d_n(\chi, \chi') \leq 2^{-k/2}$. In section §3.4.2, second moment arguments are used to reduce equation (3.3) to the following:

Proposition 3.4.11. *Let $0 < \eta_0 < (1 - \log 2)/2$. Then for all sufficiently large k (depending on η_0), if*

$$r := d/k = \frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2 + \eta$$

for some η with $\eta_0 \leq \eta < (1 - \log 2)/2$ then with high probability in the planted model, $|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)$. In symbols,

$$\lim_{n\to\infty} \mathbb{P}_n^\chi\big(|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)\big) = 1.$$

In §3.5, Proposition 3.4.11 is reduced as follows. First, certain subsets of vertices are defined through local combinatorial constraints. There are two main lemmas concerning these subsets; one of which bounds their density and the other proves they are 'rigid'. Proposition 3.4.11 is proven in §3.5 assuming these lemmas.

The density lemma is proven in §3.6 using a natural Markov model on the space of proper colorings that is the local-on-average limit of the planted model. Rigidity is proven in §3.7 using an expansivity argument similar to the way random regular graphs are proven to be good expanders. This completes the last step of the proof of Theorem 1.3.1.

3.1 Topological sofic entropy

This section defines topological sofic entropy for subshifts using the formulation from [30]. The main result is:

Lemma 3.1.1. *For any sofic approximation $\Sigma = \{\sigma_n\}$ with $\sigma_n \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$,*

$$h_\Sigma(\Gamma \curvearrowright X) = \inf_{\epsilon>0} \limsup_{n\to\infty} n^{-1} \log Z(\epsilon; \sigma_n).$$

Let Γ denote a countable group, $\mathcal{A}$ a finite set (called the **alphabet**). Let $T = (T^g)_{g\in\Gamma}$ be the shift action on $\mathcal{A}^\Gamma$ defined by $T^g x(f) = x(g^{-1}f)$ for $x \in A^\Gamma$. Let $X \subset \mathcal{A}^\Gamma$ be a closed Γ-invariant subspace. We denote the restriction of the action to X by $\Gamma \curvearrowright X$. Also let $\Sigma = \{\sigma_i : \Gamma \to \mathrm{Sym}(V_i)\}_{i\in\mathbb{N}}$ be a sofic approximation to Γ.

Given $\sigma : \Gamma \to \mathrm{Sym}(V)$, $v \in V$ and $x : V \to \mathcal{A}$ the **pullback name of x at v** is defined by

$$\Pi_v^\sigma(x) \in \mathcal{A}^\Gamma, \quad \Pi_v^\sigma(x)(g) = x_{\sigma(g^{-1})v} \quad \forall g \in \Gamma.$$

For the sake of building some intuition, note that when σ is a homomorphism, the map $v \mapsto \Pi_v^\sigma(x)$ is Γ-equivariant (in the sense that $\Pi_{\sigma(g)v}^\sigma(x) = g\Pi_v^\sigma(x)$). In particular $\Pi_v^\sigma(x) \in \mathcal{A}^\Gamma$ is periodic. In general, we think of $\Pi_v^\sigma(x)$ as an approximate periodic point.

Given an open set $\mathcal{O} \subset \mathcal{A}^\Gamma$ containing X and an $\epsilon > 0$, a map $x : V \to \mathcal{A}$ is called an **$(\mathcal{O}, \epsilon, \sigma)$-microstate** if

$$\#\{v \in V : \ \Pi_v^\sigma(x) \in \mathcal{O}\} \geq (1-\epsilon)|V|.$$

Let $\Omega(\mathcal{O}, \epsilon, \sigma) \subset \mathcal{A}^V$ denote the set of all $(\mathcal{O}, \epsilon, \sigma)$-microstates. Finally, the **Σ-entropy** of the action is defined by

$$h_\Sigma(\Gamma \curvearrowright X) := \inf_{\mathcal{O}} \inf_{\epsilon>0} \limsup_{i\to\infty} |V_i|^{-1} \log \#\Omega(\mathcal{O}, \epsilon, \sigma_i)$$

where the infimum is over all open neighborhoods of X in $\mathcal{A}^\Gamma$. This number depends on the action $\Gamma \curvearrowright X$ only up to topological conjugacy. It is an exercise in [22] to show that this definition agrees with the definition in [31]. We include a proof in Appendix A for completeness.

Proof of Lemma 3.1.1. Let $\epsilon > 0$ be given. Let $S(\epsilon; \sigma_n) \subset 2^{V_n}$ be the set of (ϵ, σ_n)-proper 2-colorings. Let $\mathcal{O}_0 \subset 2^\Gamma$ be the set of all 2-colorings $\chi : \Gamma \to \{0, 1\}$ such that for each generator hyper-edge $e \subset \Gamma$, $\chi(e) = \{0, 1\}$. A **generator hyper-edge** is a subgroup of the form $\{s_i^j : \ 0 \leq j < k\}$ for some i. Note $\mathcal{O}_0$ is an open superset of X.

We claim that $\Omega(\mathcal{O}_0, k\epsilon/d, \sigma_n) \subset S(\epsilon; \sigma_n)$. To see this, let $\chi \in \Omega(\mathcal{O}_0, k\epsilon/d, \sigma_n)$. Then $\Pi_v^{\sigma_n}(\chi) \in \mathcal{O}_0$ if and only if all hyper-edges of G_σ containing v are bi-chromatic (with respect to χ). So if $\Pi_v^{\sigma_n} \notin \mathcal{O}_0$, then v is contained in up to d monochromatic hyperedges. On the other hand, each monochromatic hyperedge contains exactly k vertices whose pullback name is not in $\mathcal{O}_0$. It follows that $\chi \in S(\epsilon; \sigma_n)$. This implies $h_\Sigma(\Gamma \curvearrowright X) \leq \inf_{\epsilon > 0} \limsup_{n \to \infty} n^{-1} \log Z(\epsilon; \sigma_n)$.

Given a finite subset $\mathcal{F}$ of hyper-edges of the Cayley hyper-tree, let $\mathcal{O}_\mathcal{F}$ be the set of all $\chi \in 2^\Gamma$ with the property that $\chi(e) = \{0, 1\}$ for all $e \in \mathcal{F}$. If $\mathcal{O}'$ is any open neighborhood of X in 2^Γ then $\mathcal{O}'$ contains $\mathcal{O}_\mathcal{F}$ for some $\mathcal{F}$. To see this, suppose that there exist elements $\chi_\mathcal{F} \in \mathcal{O}_\mathcal{F} \setminus \mathcal{O}'$ for every finite $\mathcal{F}$. Let χ be a cluster point of $\{\chi_\mathcal{F}\}$ as $\mathcal{F}$ increases to the set E of all hyper-edges. Then $\chi \in X \setminus \mathcal{O}'$, a contradiction. It follows that

$$h_\Sigma(\Gamma \curvearrowright X) = \inf_\mathcal{F} \inf_{\epsilon > 0} \limsup_{i \to \infty} |V_i|^{-1} \log \#\Omega(\mathcal{O}_\mathcal{F}, \epsilon, \sigma_i).$$

Next, fix a finite subset $\mathcal{F}$ of hyper-edges of the Cayley hyper-tree. We claim that $S\left(\frac{\epsilon}{k|\mathcal{F}|}; \sigma_n\right) \subset \Omega(\mathcal{O}_\mathcal{F}, \epsilon, \sigma_n)$. To see this, let $\chi \in S\left(\frac{\epsilon}{k|\mathcal{F}|}; \sigma_n\right)$ and $B(\chi, \sigma_n) \subset V_n$ be the set of vertices contained in a monochromatic edge of χ. Now for $v \in V_n$, $\Pi_v^{\sigma_n}(\chi) \notin \mathcal{O}_\mathcal{F}$ if and only if $\Pi_v^{\sigma_n}(\chi)$ is monochromatic on some edge in $\mathcal{F}$. This occurs if and only if there is an element $f \in \Gamma$ in the union of $\mathcal{F}$ such that $\sigma_n(f^{-1})v \in B(\chi, \sigma_n)$. There are

at most $k|\mathcal{F}||B(\chi,\sigma_n)|$ such vertices. But $|B(\chi,\sigma_n)| \leq (\frac{\epsilon}{|\mathcal{F}|})n$, so there are at most $k\epsilon n$ such vertices. It follows that $\chi \in \Omega(\mathcal{O}_{\mathcal{F}},\epsilon,\sigma_n)$. Therefore,

$$\inf_{\epsilon>0}\limsup_{n\to\infty} n^{-1}\log Z(\epsilon;\sigma_n) \leq \inf_{\mathcal{F}}\inf_{\epsilon>0}\limsup_{i\to\infty}|V_i|^{-1}\log\#\Omega(\mathcal{O}_{\mathcal{F}},\epsilon,\sigma_i) = h_\Sigma(\Gamma\curvearrowright X).$$

□

3.2 Reduction to the key lemma

The purpose of this section is to show how Lemma 3.0.6 implies Theorem 1.3.1. This requires replacing the (random) uniform and planted models with (deterministic) sofic approximations. The next lemma facilitates this replacement.

Lemma 3.2.1. *Let $D \subset \Gamma$ be finite and $\delta > 0$. Then there are constants $\epsilon, N_0 > 0$ such that for all $n > N_0$ with $n \in 2\mathbb{Z} \cap k\mathbb{Z}$,*

$$\mathbb{P}_n^u\{\sigma :\ \sigma \text{ is not } (\mathrm{D},\delta)\text{−sofic}\} \leq \mathrm{n}^{-\epsilon\mathrm{n}},$$

$$\mathbb{P}_n^p\{\sigma :\ \sigma \text{ is not } (\mathrm{D},\delta)\text{−sofic}\} \leq \mathrm{n}^{-\epsilon\mathrm{n}}.$$

Proof. The proof given here is for the uniform model. The planted model is similar.

The proof begins with a series of four reductions. By taking a union bound, it suffices to prove the special case in which $D = \{w\}$ for $w \in \Gamma$ nontrivial. (This is the first reduction).

Let $w = s_{i_l}^{r_l}\cdots s_{i_1}^{r_1}$ be the **reduced form** of w. This means that $i_j \in \{1,\ldots,d\}$, $i_j \neq i_{j+1}$ for all j with indices mod l and $1 \leq r_j < k$ for all j. Let $|w| = r_1 + \cdots + r_l$ be the **length** of w.

For any $g \in \Gamma$, the fixed point sets of $\sigma(gwg^{-1})$ and $\sigma(w)$ have the same size. So after conjugating if necessary, we may assume that either $l = 1$ or $i_1 \neq i_l$.

For $1 \leq j \leq l$, the **j-th beginning subword** of w is the element $w_j = s_{i_j}^{r_j} \cdots s_{i_1}^{r_1}$. Given a vertex $v \in V_n$ and $\sigma \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$, let $p(v, \sigma) = (e_1, \ldots, e_l)$ be the path defined by: for each j, e_j is the unique hyper-edge of G_σ labeled i_j containing $\sigma(w_j)v$. A vertex $v \in V_n$ **represents a (σ, w)-simple cycle** if $\sigma(w)v = v$ and for every $1 \leq a < b \leq l$, either

- $e_a \cap e_b = \emptyset$,
- $b = a + 1$ and $|e_a \cap e_b| = 1$,
- or $(a, b) = (1, l)$ and $|e_a \cap e_b| = 1$.

We say that v **represents a (σ, w)-simple degenerate cycle** if $\sigma(w)v = v$ and $l = 2$ and $|e_1 \cap e_2| \geq 2$.

If $\sigma(w)v = v$ then either

- v represents a (σ, w)-simple cycle,
- there exists nontrivial $w' \in \Gamma$ with $|w'| \leq |w| + k$ such that some vertex $v_0 \in \cup_j e_j$ represents a (σ, w')-simple cycle,
- or there exists nontrivial $w' \in \Gamma$ with $|w'| \leq |w| + k$ such that some vertex $v_0 \in \cup_j e_j$ represents a (σ, w')-simple degenerate cycle.

So it suffices to prove there are constants $\epsilon, N_0 > 0$ such that for all $n > N_0$,

$$\mathbb{P}_n^u\big\{\sigma : \ \#\{v \in [n] : \ v \text{ represents a } (\sigma, w)\text{-simple cycle}\} \geq \delta n\big\} \leq n^{-\epsilon n}.$$

and

$$\mathbb{P}_n^u\big\{\sigma : \ \#\{v \in [n] : \ v \text{ represents a } (\sigma, w)\text{-simple degenerate cycle}\} \geq \delta n\big\} \leq n^{-\epsilon n}.$$

(This is the second reduction).

Two vertices $v, v' \in V_n$ **represent vertex-disjoint** (σ, w)**-cycles** if $p(v,\sigma) = (e_1, \ldots, e_l), p(v',\sigma) = (e'_1, \ldots, e'_l)$ and $e_i \cap e'_j = \emptyset$ for all i, j.

Let $G_n(\delta, w)$ be the set of all $\sigma \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$ such that there exists a subset $S \subset [n]$ satisfying

1. $|S| \geq \delta n$,

2. every $v \in S$ represents a (σ, w)-simple cycle,

3. the cycles $p(v,\sigma)$ for $v \in S$ are pairwise vertex-disjoint.

If v represents a simple (σ, w)-cycle then there are at most $(kl)^2$ vertices v' such that v' also represents a simple (σ, w)-cycle but the two cycles are not vertex-disjoint. Since this bound does not depend on n, it suffices to prove there exist $\epsilon > 0$ and N_0 such that

$$\mathbb{P}_n^u(G_n(\delta, w)) \leq n^{-\epsilon n}$$

for all $n \geq N_0$. (This is the third reduction. The argument is similar for simple degenerate cycles).

Let $m = \lceil \delta n \rceil$ and $v_1, \ldots, v_m$ be distinct vertices in $[n] = V_n$. For $1 \leq i \leq m$, let F_i be the set of all $\sigma \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$ such that for all $1 \leq j \leq i$

1. v_j represents a (σ, w)-simple cycle,

2. the cycles $p(v_1,\sigma), \ldots, p(v_i,\sigma)$ are pairwise vertex-disjoint.

By summing over all subsets of size m, we obtain

$$\mathbb{P}_n^u(G_n(\delta, w)) \leq \binom{n}{m} \mathbb{P}_n^u(F_m).$$

Since $\binom{n}{m} \approx e^{H(\delta, 1-\delta)n}$ grows at most exponentially, it suffices to show there exist $\epsilon > 0$ and N_0 such that $\mathbb{P}_n^u(F_m) \leq n^{-\epsilon n}$ for all $n \geq N_0$. (This is the fourth reduction. The argument is similar for simple degenerate cycles).

Set $F_0 = \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$. By the chain rule

$$\mathbb{P}_n^u(F_m) = \prod_{i=0}^{m-1} \mathbb{P}_n^u(F_{i+1}|F_i).$$

In order to estimate $\mathbb{P}_n^u(F_{i+1}|F_i)$, F_i can be expressed a disjoint union over the cycles involved in its definition. To be precise, define an equivalence relation $\mathcal{R}_i$ on F_i by: σ, σ' are $\mathcal{R}_i$-equivalent if for every $1 \leq j \leq i$, $1 \leq q \leq l$ and $r > 0$

$$\sigma(s_{i_q}^r w_q)v_j = \sigma'(s_{i_q}^r w_q)v_j.$$

In other words, σ, σ' are $\mathcal{R}_i$-equivalent if they define the same paths according to all vertices up to v_i (so $p(v_j, \sigma) = p(v_j, \sigma')$) and their restrictions to every edge in these paths agree. Of course, F_i is the disjoint union of the $\mathcal{R}_i$-classes. Note that $\mathcal{R}_0$ is trivial (everything is equivalent).

In general, if $A, B_1, \ldots, B_m$ are measurable sets and the B_i's are pairwise disjoint then $\mathbb{P}(A|\cup_i B_i)$ is a convex combination of $\mathbb{P}(A|B_i)$ (for any probability measure $\mathbb{P}$). Therefore, $\mathbb{P}_n^u(F_{i+1}|F_i)$ is a convex combination of probabilities of the form $\mathbb{P}_n^u(F_{i+1}|B_i)$ where B_i is an $\mathcal{R}_i$-class.

Now fix a $\mathcal{R}_i$-class B_i (for some i with $0 \leq i < m$). Let K be the set of all vertices covered by the cycles defining B_i. To be precise, this means K is the set of all $u \in [n] = V_n$ such that there exists an edge e with $u \in e$ such that e is contained in a path $p(v_j, \sigma)$ with $1 \leq j \leq i$ and $\sigma \in B_i$. Since each path covers at most kl vertices, $|K| \leq ikl$.

If $l > 1$ (the case $l = 1$ is similar), fix subsets $e_1, \ldots, e_{l-1} \subset [n]$ of size k. Conditioned on B_i and the event that the first $(l-1)$ edges of $p(v_{i+1}, \sigma)$ are $e_1, \ldots, e_{l-1}$, the $\mathbb{P}_n^u$-probability that v_{i+1} represents a simple (σ, w)-cycle vertex-disjoint from K is

bounded by the probability that a uniformly random $k-1$-element subset of

$$[n] \setminus \left(\bigcup_{2 \leq j \leq l-1} e_j \cup K \right)$$

contains v_{i+1}. Since

$$\left| \bigcup_{2 \leq j \leq l-1} e_j \cup K \right| \leq (i+1)kl \leq mkl = kl\lceil \delta n \rceil,$$

this probability is bounded by C/n where $C = C(w, d, k, \delta)$ is a constant not depending on n or the choice of B_i. It follows that $\mathbb{P}_n^u(F_{i+1}|F_i) \leq C/n$ for all $0 \leq i \leq m-1$ and therefore

$$\mathbb{P}_n^u(F_m) \leq (C/n)^m \leq (C/n)^{\delta n}.$$

This implies the lemma (the argument is similar for simple degenerate cycles).

□

Proof of Theorem 1.3.1 from Lemma 3.0.6. Choose $d, k, \eta_0, \eta_1, \eta$ according to the hypotheses of Lemma 3.0.6 so that $\eta \in [\eta_0, \eta_1]$, $k \geq k_0$ and all of the conclusions to Lemma 3.0.6 hold. Construct $\Gamma \curvearrowright X$ according to Section 1.3. We will always take $n \in 2\mathbb{Z} \cap k\mathbb{Z}$. We will first show the existence of a sofic approximation $\Sigma_1 = \{\sigma_n\}$ to Γ such that

$$\inf_{\epsilon > 0} \limsup_{n \to \infty} n^{-1} \log Z(\epsilon; \sigma_n) = f(d, k).$$

Then by Lemma 3.1.1, we will have $h_{\Sigma_1}(\Gamma \curvearrowright X) = f(d, k) > 0$.

Observe that Lemma 3.0.6(1) implies the following: for every $\beta > 0$ there exists $\epsilon(\beta) > 0$ (with ϵ decreasing as β decreases), N, and $\xi(\beta, \epsilon, N) > 0$ such that for $n > N$, $\mathbb{P}_n^u(n^{-1} \log(Z(\epsilon; \sigma_n)) - f(d, k) \geq \beta) < e^{-\xi n}$. In other words, the number of ϵ-proper colorings can only exponentially exceed $f(d, k)$ with exponentially small probability. This is because the negation of the above would imply that

$\inf_{\epsilon>0} \limsup_{n\to\infty} n^{-1} \log \mathbb{E}_n^u[Z(\epsilon;\sigma)] > f(d,k)$, contradicting (1). Let $F_{n,\beta,\epsilon}$ be the event that $n^{-1}\log[Z(\epsilon;\sigma_n)] - f(d,k) \geq \beta$.

Let $H_{n,\beta}$ be the event that $|n^{-1}\log Z(\sigma_n) - f(d,k)| < \beta$. Lemma 3.0.6(3) implies $\mathbb{P}_n^u(H_{n,\beta})$ decays at most subexponentially in n. Precisely, for any $c > 0$ there exists $N = N(c,\beta)$ such that $n > N$ implies $\mathbb{P}_n^u(H_{n,\beta}) > e^{-cn}$.

Let $I_{n,\beta,\epsilon}$ be the event that $|n^{-1}\log Z(\epsilon;\sigma_n) - f(d,k)| < \beta$. Notice that since $Z(\epsilon;\sigma_n) \geq Z(\sigma_n)$ for any ϵ, $F_{n,\beta,\epsilon}^c \cap H_{n,\beta} \subset I_{n,\beta,\epsilon}$.

Consider a decreasing sequence $\beta_m \to 0$, and ϵ_m and ξ_m depending on β_m as discussed above. Since $\mathbb{P}_n^u(H_{n,\beta})$ is decaying only subexponentially, we can choose an increasing sequence K_m satisfying $K_m \geq \max_{1\leq j\leq m} N_j$, and for each m, $\mathbb{P}_n^u(H_{n,\beta_m}) > 2\sum_{i=1}^m e^{-\xi_i n}$ for all $n > K_m$. Now it follows that $n > K_m$ implies $\mathbb{P}_n^u(\cap_{j=1}^m F_{n,\beta_j,\epsilon_j}^c \cap H_{n,\beta_j}) \geq \mathbb{P}_n^u(H_{n,\beta_m}) - \sum_{i=1}^m e^{-\xi_i n} \geq e^{-\xi_1 n}$. By the observation in the preceding paragraph, $\mathbb{P}_n^u(\cap_{j=1}^m I_{n,\beta_j,\epsilon_j}) \geq e^{-\xi_1 n}$.

Given $\delta > 0$ and $D \subset \Gamma$ finite, let $J_{D,\delta,n}$ be the event that σ_n is (D,δ)-sofic. Let $\delta_m \to 0$ be a decreasing sequence and $D_m \subset \Gamma$ be an increasing sequence of finite subsets of Γ.

Lemma 3.2.1 implies that $\mathbb{P}_n^u(J_{D,\delta,n}^c)$ decays super-exponentially in n for any D and δ, so there exists an increasing sequence N_m such that for $n > N_m$, $\mathbb{P}_n^u(\cap_{j=1}^m I_{n,\beta_j,\epsilon_j} \cap J_{D_m,\delta_m,n}) \geq 0.5e^{-\xi_1 n}$. It follows that we can choose a deterministic sofic approximation sequence $\Sigma_1 = \{\sigma_n\}$ such that

$$\inf_{\epsilon>0} \limsup_{n\to\infty} n^{-1}\log Z(\epsilon;\sigma_n) = f(d,k).$$

We next show the existence of Σ_2. Equation (2) of Lemma 3.0.6 implies the existence of a number f_p with

$$f(d,k) < f_p < \liminf_{n\to\infty} n^{-1}\log \mathbb{E}_n^p[Z(\sigma)].$$

Since $Z(\sigma) \leq 2^n$ for every σ, there exist constants $c, N_0 > 0$ such that

$$\mathbb{P}_n^p\{\sigma : \ Z(\sigma) \geq \exp(nf_p)\} \geq \exp(-cn) \tag{3.4}$$

for all $n \geq N_0$.

Now let $\delta > 0$ and $D \subset \Gamma$ be finite. Then there exists N_2 such that if $n > N_2$ and σ_n is chosen at random with law $\mathbb{P}_n^p$, then with positive probability,

1. σ_n is (D, δ)-sofic,

2. $n^{-1} \log Z(\sigma_n) \geq f_p$.

This is implied by Lemma 3.2.1 and equation (3.4). So there exists a sofic approximation $\Sigma_2 = \{\sigma_n'\}$ to Γ such that

$$\limsup_{n\to\infty} n^{-1} \log Z(\sigma_n') \geq f_p.$$

Since $Z(\sigma_n') \leq Z(\beta; \sigma_n')$, Lemma 3.1.1 implies $h_{\Sigma_2}(\Gamma \curvearrowright X) \geq f_p > f(d,k) = h_{\Sigma_1}(\Gamma \curvearrowright X)$.

□

3.3 The first moment

To simplify notation, we assume throughout the paper that $n \in 2\mathbb{Z} \cap k\mathbb{Z}$ without further mention. This section proves (3.1) of Lemma 3.0.6. The proof is in two parts. Part 1, in §3.3.1, establishes:

Theorem 3.3.1.

$$\lim_{\epsilon \searrow 0} \limsup_{n\to\infty} (1/n) \log \mathbb{E}_n^u[Z(\epsilon; \sigma)] = \limsup_{n\to\infty} (1/n) \log \mathbb{E}_n^u[Z(\sigma)].$$

Part 2 has to do with equitable colorings, where a 2-coloring $\chi : [n] \to \{0, 1\}$ is

equitable if

$$|\chi^{-1}(0)| = |\chi^{-1}(1)| = n/2.$$

Let $Z_e(\sigma)$ be the number of proper equitable colorings of G_σ. §3.3.2 establishes

Theorem 3.3.2.

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_n^u[Z(\sigma)] = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_n^u[Z_e(\sigma)].$$

Moreover,

$$\frac{1}{n} \log \mathbb{E}_n^u[Z_e(\sigma)] = f(d,k) + O(n^{-1}\log(n))$$

where $f(d,k) = \log(2) + \frac{d}{k}\log(1-2^{1-k})$.

Combined, Theorems 3.3.1 and 3.3.2 imply (3.1) of Lemma 3.0.6.

Remark 3.3.3. *If* $r := (d/k)$ *then the formula for* $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_n^u[Z(\sigma)]$ *above is the same as the formula found in [27, 28, 29] for the exponential growth rate of the number of proper 2-colorings of* $H_k(n,m)$.

Remark 3.3.4. *When we write an error term, such as* $O(n^{-1}\log(n))$, *we always assume that* $n \geq 2$ *and the implicit constant is allowed to depend on* k *or* d.

3.3.1 Almost proper 2-colorings

For $0 < x \leq 1$, let $\eta(x) = -x\log(x)$. Also let $\eta(0) = 0$. If $\vec{T} = (T_i)_{i\in I}$ is a collection of numbers with $0 \leq T_i \leq 1$, then let

$$H(\vec{T}) := \sum_{i\in I} \eta(T_i)$$

be the **Shannon entropy** of $\vec{T}$.

Definition 3.3.5. *A* k**-partition** *of* $[n]$ *is an unordered partition of* $[n]$ *into sets of size* k*. Of course, such a partition exists if and only if* $n/k \in \mathbb{N}$ *in which case there are*

$$\frac{n!}{k!^{n/k}(n/k)!} \tag{3.5}$$

such partitions. By Stirling's formula,

$$\frac{1}{n}\log(\#\{k\text{-partitions}\}) = (1-1/k)(\log(n)-1) - (1/k)\log(k-1)! + O(n^{-1}\log(n)) \tag{3.6}$$

Definition 3.3.6. *The* **orbit-partition** *of a permutation* $\rho \in \mathrm{Sym}(n)$ *is the partition of* $[n]$ *into orbits of* ρ*. Fix a* k*-partition* π*. Then the number of permutations* ρ *whose orbit partition is* π *equals* $(k-1)!^{n/k}$*.*

Given $\sigma \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$*, define the* d*-tuple* $(\pi_1^\sigma, \ldots, \pi_d^\sigma)$ *of* k*-partitions by:* π_i^σ *is the orbit-partition of* $\sigma(s_i)$*. Fix a* d*-tuple of* k*-partitions* $(\pi_1, \ldots, \pi_d)$*. Then the number of uniform homomorphisms* σ *such that* $\pi_i^\sigma = \pi_i$ *for all* i *is* $[(k-1)!^{n/k}]^d$*. Combined with (3.5), this shows the number of uniform homomorphisms into* $\mathrm{Sym}(n)$ *is*

$$\left[\frac{n!(k-1)!^{n/k}}{k!^{n/k}(n/k)!}\right]^d.$$

By Stirling's formula,

$$\frac{1}{n}\log \#\mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n)) = d(1-1/k)(\log n - 1) + O(n^{-1}\log(n)). \tag{3.7}$$

Definition 3.3.7. *Let* π *be a* k*-partition,* $\chi : [n] \to \{0,1\}$ *a 2-coloring and* $\vec{t} = (t_j)_{j=0}^k \in [0,1]^{k+1}$ *a vector with* $\sum_j t_j = 1/k$*. The pair* (π, χ) *has* **type** $\vec{t}$ *if for all* j*,*

$$\#\left\{e \in \pi : \ |e \cap \chi^{-1}(1)| = j\right\} = nt_j.$$

Lemma 3.3.8. *Let* $\vec{t} = (t_0, t_1, \ldots, t_k) \in [0,1]^{k+1}$ *be such that* $\sum_j t_j = 1/k$ *and* $nt_j \in \mathbb{Z}$. *Let* $p = \sum_j jt_j$. *Let* $\chi : [n] \to \{0,1\}$ *be a map such that* $|\chi^{-1}(1)| = pn$. *Let* $f(\vec{t})$ *be the number of* k*-partitions* π *of* $[n]$ *such that* (π, χ) *has type* $\vec{t}$. *Then*

$$(1/n)\log f(\vec{t}) = (1-1/k)(\log(n)-1) - H(p,1-p) + H(\vec{t}) - \sum_{j=0}^{k} t_j \log(j!(k-j)!) + O(n^{-1}\log(n)).$$

Proof. The following algorithm constructs all such partitions with no duplications:

Step 1. Choose an unordered partition of the set $\chi^{-1}(1)$ into $t_j n$ sets of size j ($j = 0, \ldots, k$).

Step 2. Choose an unordered partition of the set $\chi^{-1}(0)$ into $t_j n$ sets of size $k - j$ ($j = 0, \ldots, k$).

Step 3. Choose a bijection between the collection of subsets of size j constructed in part 1 with the collection of subsets of size $k - j$ constructed in part 2.

Step 4. The partition consists of all sets of the form $\alpha \cup \beta$ where $\alpha \subset \chi^{-1}(1)$ is a set of size j constructed in Step 1 and $\beta \subset \chi^{-1}(0)$ is a set of size $(k-j)$ constructed in Step 2 that it is paired with under Step 3.

The number of choices in Step 1 is $\frac{(pn)!}{\prod_{j=1}^{k}(j)!^{t_j n}(t_j n)!}$. The number of choices in Step 2 is $\frac{((1-p)n)!}{\prod_{j=0}^{k-1}(k-j)!^{t_j n}(t_j n)!}$. The number of choices in Step 3 is $\prod_{j=1}^{k-1}(t_j n)!$. So

$$f(\vec{t}) = \frac{(pn)!((1-p)n)!}{\prod_{j=0}^{k} j!^{t_j n}(k-j)!^{t_j n}(t_j n)!}. \tag{3.8}$$

The lemma follows from this and Stirling's formula. □

Let $\mathcal{M}$ be the set of all matrices $\vec{T} = (T_{ij})_{1 \le i \le d, 0 \le j \le k}$ such that

1. $T_{ij} \ge 0$ for all i, j,
2. $\sum_{j=0}^{k} T_{ij} = 1/k$ for all i,

3. there exists a number, denoted $p(\vec{T})$, such that $p(\vec{T}) = \sum_{j=0}^{k} jT_{ij}$ for all i.

Let $\mathcal{M}_n$ be the set of all $\vec{T} \in \mathcal{M}$ such that $n\vec{T}$ is integer-valued.

Lemma 3.3.9. *Let A be an integer-valued $q \times p$ matrix (for some $p, q \in \mathbb{N}$), $b \in \mathbb{R}^q$ and $\mathcal{K} \subset \mathbb{R}^p$ the set of all $x \in \mathbb{R}^p$ such that $Ax = b$ and $x_i \geq 0$ for all i. For $n \in \mathbb{N}$, let $\mathcal{K}_n$ be the set of all $x \in \mathcal{K}$ such that nx is integer-valued.*

Assume $\mathcal{K}$ is compact and there exists $x \in \mathcal{K}$ with $x_i > 0$ for all i. Then there is a constant $C > 0$ such that if $\mathcal{K}_n$ is non-empty then it is C/n-dense in $\mathcal{K}$ in the following sense. For any $x \in \mathcal{K}$ there exists $x' \in \mathcal{K}_n$ such that $\|x - x'\|_\infty < C/n$. Moreover, the constant C can be chosen to depend continuously on the vector $b \in \mathbb{R}^q$.

Proof. To begin, we will define several constants which will enable us to choose $C > 0$. Because A is integer-valued, its kernel, denoted $\ker(A) \subset \mathbb{R}^p$, is such that $\ker(A) \cap \mathbb{Z}^p$ has rank equal to the dimension of $\ker(A)$. Therefore, $\ker(A) \cap \mathbb{Z}^p$ is cocompact in $\ker(A)$. So there is a constant $C_1 > 0$ such that for any $z \in \ker(A)$ there is an element $z' \in \ker(A) \cap \mathbb{Z}^p$ with $\|z - z'\|_\infty < C_1$.

By hypothesis, there is a constant $C_2 > 0$ and an element $y \in \mathcal{K}$ such that $y_i > C_2$ for all i. Because $\mathcal{K}$ is compact there is another constant $C_3 > 0$ such that $\|x - y\|_\infty < C_3$ for all $x \in \mathcal{K}$. Let $C = \frac{C_1 C_3}{C_2} + C_1$. Now let $x \in \mathcal{K}$ be arbitrary and suppose $\mathcal{K}_n$ is non-empty. We will show there exists $x'' \in \mathcal{K}_n$ such that $\|x - x''\|_\infty \leq C/n$.

Let

$$x' = \left(1 - \frac{C_1}{C_2 n}\right) x + \frac{C_1}{C_2 n} y.$$

Then $x'_i \geq (C_1/C_2 n) y_i > C_1/n$ for all i. Also $\|x - x'\|_\infty = (C_1/C_2 n)\|x - y\|_\infty < \frac{C_1 C_3}{C_2 n}$.

Because $\mathcal{K}_n$ is non-empty, there exists $x_n \in \mathcal{K}_n$. By linearity, $\mathcal{K}$ is the intersection of the hyperplane $x_n + \ker(A)$ with the positive orthant. Thus we can write $x' = x_n + (1/n)z' + z''$ where $z' \in \ker(A) \cap \mathbb{Z}^p$ and $z'' \in \ker(A)$ satisfies $\|z''\|_\infty < C_1/n$. Let $x'' = x_n + (1/n)z'$. Note $\|x'' - x'\|_\infty = \|z''\|_\infty < C_1/n$. Since $x'_i > C_1/n$ this implies

$x''_i > 0$ for all i. It is now straightforward to check that $x'' \in \mathcal{K}_n$. By the triangle inequality

$$\|x - x''\|_\infty \leq \|x - x'\|_\infty + \|x' - x''\|_\infty < \frac{C_1 C_3}{C_2 n} + C_1/n = C/n.$$

Because x and n are arbitrary, this implies the Lemma with $C = \frac{C_1 C_3}{C_2} + C_1$. Moreover C_1 does not depend on the vector b; while C_2, C_3 can be chosen to depend continuously on b.

□

Lemma 3.3.10. *Given a matrix $\vec{T} \in \mathcal{M}$ define*

$$F(\vec{T}) \;:=\; H(\vec{T}) + (1-d)H(p, 1-p) - (d/k)\log k + \sum_{i=1}^{d}\sum_{j=0}^{k} T_{ij} \log \binom{k}{j}$$

where $p = p(\vec{T})$. Then for any $\epsilon \geq 0$,

$$(1/n)\log \mathbb{E}_n^u[Z(\epsilon;\sigma)] = \sup\left\{ F(\vec{T}) :\ \vec{T} \in \mathcal{M}_n \text{ and } \sum_{i=1}^{d}\sum_{j=0,k} T_{ij} \leq \epsilon \right\} + O(n^{-1}\log(n))$$

where the constant implicit in the error term does not depend on ϵ.

Proof. Given $\sigma \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$ and $1 \leq i \leq d$, let π_i^σ be the orbit-partition of $\sigma(s_i)$. For $\vec{T}$ as above, let $Z_\sigma(\vec{T})$ be the number of ϵ-proper colorings $\chi : [n] \to \{0,1\}$ such that (π_i^σ, χ) has type $\vec{T}_i = (T_{i,0}, \ldots, T_{i,k})$. It suffices to show that

$$(1/n)\log \mathbb{E}_n^u[Z_\sigma(\vec{T})] = F(\vec{T}) + O(n^{-1}\log(n))$$

for all $n \geq 2$ such that $\vec{T} \in \mathcal{M}_n$. This is because the size of $\mathcal{M}_n$ is a polynomial (depending on k, d) in n so the supremum above determines the exponential growth rate of $\mathbb{E}_n^u[Z(\epsilon;\sigma)]$.

To prove this, fix a $\vec{T}$ as above and let n be such that $n\vec{T}$ is integer-valued. Fix a

coloring $\chi : [n] \to \{0,1\}$ such that $|\chi^{-1}(1)| = pn$. By symmetry,

$$\mathbb{E}_n^u[Z_\sigma(\vec{T})] = \binom{n}{pn} \mathbb{P}_n^u[(\pi_i^\sigma, \chi) \text{ has type } \vec{T}_i \ \forall i].$$

The events $\{(\pi_i^\sigma, \chi) \text{ has type } \vec{T}_i\}_{i=1}^d$ are jointly independent. So

$$\mathbb{E}_n^u[Z_\sigma(\vec{T})] = \binom{n}{pn} \prod_{i=1}^d \mathbb{P}_n^u[(\pi_i^\sigma, \chi) \text{ has type } \vec{T}_i]. \tag{3.9}$$

By symmetry, $\mathbb{P}_n^u[(\pi_i^\sigma, \chi) \text{ has type } \vec{T}_i]$ is the number of k-partitions π such that (π, χ) has type $\vec{T}_i$ divided by the number of k-partitions. By Lemma 3.3.8 and (3.6),

$$\begin{aligned} & \frac{1}{n} \log \mathbb{P}_n^u[(\pi_i^\sigma, \chi) \text{ has type } \vec{T}_i] \\ = \ & -H(p, 1-p) + H(\vec{T}_i) - \sum_{j=0}^k T_{ij} \log(j!(k-j)!) + (1/k) \log(k-1)! + O(n^{-1} \log(n)). \end{aligned}$$

Combine this with (3.9) to obtain

$$\begin{aligned} & (1/n) \log \mathbb{E}_n^u[Z_\sigma(\vec{T})] \\ = \ & (1-d)H(p, 1-p) + H(\vec{T}) - \sum_{i=1}^d \sum_{j=0}^k T_{ij} \log(j!(k-j)!) + (d/k) \log(k-1)! + O(n^{-1} \log(n)). \end{aligned}$$

This simplifies to the formula for $F(\vec{T})$ using the assumption that $\sum_{j=0}^k T_{ij} = 1/k$ for all i.

□

Proof of Theorem 3.3.1. By Lemma 3.3.9 applied to $\mathcal{M}$, continuity of F and com-

pactness of $\mathcal{M}$,

$$\lim_{n\to\infty} \sup \left\{ F(\vec{T}) : \ \vec{T} \in \mathcal{M}_n \text{ and } \sum_{i=1}^{d} \sum_{j=0,k} T_{ij} \leq \epsilon \right\}$$
$$= \sup \left\{ F(\vec{T}) : \ \vec{T} \in \mathcal{M} \text{ and } \sum_{i=1}^{d} \sum_{j=0,k} T_{ij} \leq \epsilon \right\}.$$

Theorem 3.3.1 now follows from Lemma 3.3.10 by continuity of F and compactness of $\mathcal{M}$. □

3.3.2 Equitable colorings

Proof of Theorem 3.3.2. Let $\mathcal{M}_0$ be the set of all $\vec{T} \in \mathcal{M}$ such that $T_{ij} = 0$ whenever $j \in \{0, k\}$. By Lemma 3.3.10, it suffices to show that F admits a unique global maximum on $\mathcal{M}_0$ and moreover if $\vec{T} \in \mathcal{M}_0$ is the global maximum then $p(\vec{T}) = 1/2$ and $F(\vec{T}) = f(d, k)$.

The function F is symmetric in the index i. To exploit this, let $\mathcal{M}'$ be the set of all vectors $\vec{t} = (t_j)_{j=1}^{k-1}$ such that $t_j \geq 0$ for all j and $\sum_{j=1}^{k-1} t_j = 1/k$. Let

$$\begin{aligned} p(\vec{t}) &= \sum_{j=1}^{k-1} j t_j \\ F(\vec{t}) &= dH(\vec{t}) + (1-d)H(p, 1-p) - (d/k)\log k + d\sum_{j=1}^{k-1} t_j \log \binom{k}{j}. \end{aligned}$$

Note that $F(\vec{t}) = F(\vec{T})$ if $\vec{T}$ is defined by $\vec{T}_{ij} = \vec{t}_j$ for all i, j. Moreover, since Shannon entropy is strictly concave, for any $\vec{T} \in \mathcal{M}_0$, if $\vec{t}$ is defined to be the average: $\vec{t}_j = d^{-1} \sum_{i=1}^{d} \vec{T}_{ij}$ then $F(\vec{t}) \geq F(\vec{T})$ with equality if and only if $\vec{t}_j = \vec{T}_{ij}$ for all i, j. So it suffices to show that F admits a unique global maximum on $\mathcal{M}'$ and moreover if $\vec{t} \in \mathcal{M}'$ is the global maximum then $p(\vec{t}) = 1/2$ and $F(\vec{t}) = f(d, k)$.

Because $\frac{\partial H(\vec{t})}{\partial t_j} = -[\log(t_j) + 1]$, $\frac{\partial p}{\partial t_j} = j$, and $\frac{\partial H(p,1-p)}{\partial t_j} = j \log\left(\frac{1-p}{p}\right)$,

$$\frac{\partial F}{\partial t_j} = -d[\log(t_j) + 1] + (1-d)j \log\left(\frac{1-p}{p}\right) + d \log\binom{k}{j}.$$

Since this is positive infinity whenever $t_j = 0$, it follows that every maximum of F occurs in the interior of $\mathcal{M}'$. The method of Lagrange multipliers implies that, at a critical point, there exists $\lambda \in \mathbb{R}$ such that

$$\nabla F = \lambda \nabla \left(\vec{t} \mapsto \sum_j t_j\right) = (\lambda, \lambda, \ldots, \lambda).$$

So at a critical point,

$$\frac{\partial F}{\partial t_j} = -d[\log(t_j) + 1] + (1-d)j \log\left(\frac{1-p}{p}\right) + d \log\binom{k}{j} = \lambda.$$

Solve for t_j to obtain

$$t_j = \exp(-\lambda/d - 1)\binom{k}{j}\left(\frac{1-p}{p}\right)^{j(1-d)/d}.$$

Note

$$1 = k \sum_{j=1}^{k-1} t_j$$

$$1 = 1/p \sum_{j=1}^{k-1} j t_j$$

implies

$$0 = \sum_{j=1}^{k-1} (k - j/p) t_j = \sum_{j=1}^{k-1} (pk - j)\binom{k}{j}\left(\frac{1-p}{p}\right)^{j(1-d)/d}.$$

So define

$$g(x) := \sum_{j=1}^{k-1} (kx - j)\binom{k}{j}\left(\frac{1-x}{x}\right)^{j(1-d)/d}.$$

It follows from the above that $g(p(\vec{t})) = 0$ whenever $\vec{t}$ is a critical point.

We claim that $g(x) = 0$ if and only if $x = 1/2$ (for $x \in (0,1)$). The change of variables $j \mapsto k - j$ in the formula for g shows that $g(1-x) = -\left(\frac{x}{1-x}\right)^{k(1-d)/d} g(x)$. So it is enough to prove that $g(x) < 0$ for $x \in (0, 1/2)$.

To obtain a simpler formula for g, set $y(x) = \left(\frac{1-x}{x}\right)^{(1-d)/d}$. The binomial formula implies

$$\begin{aligned} g(x) &= \sum_{j=1}^{k-1} (kx - j)\binom{k}{j} y^j \\ &= kx[(1+y)^k - 1 - y^k] - ky[(1+y)^{k-1} - y^{k-1}] \\ &= k[(x(1+y) - y)(1+y)^{k-1} - x + (-x+1)y^k]. \end{aligned}$$

Because $0 < x < 1/2$, $y > \left(\frac{x}{1-x}\right)$ which implies that the middle coefficient $(x(1+y) - y) = x - y(1-x) < 0$. So

$$g(x)/k < (1-x)y^k - x < 0$$

where the last inequality holds because

$$y^k = \left(\frac{x}{1-x}\right)^{k(d-1)/d} < \frac{x}{1-x}$$

assuming $k(d-1)/d > 1$. This proves the claim.

So if $\vec{t}$ is a critical point then $p(\vec{t}) = 1/2$. Put this into the equation above for t_j to obtain

$$t_j = C\binom{k}{j}$$

where $C = \exp(-\lambda/d - 1)$. Because

$$1/k = \sum_{j=1}^{k-1} t_j = C \sum_{j=1}^{k-1} \binom{k}{j} = C(2^k - 2)$$

it must be that

$$t_j = \frac{1}{k(2^k - 2)} \binom{k}{j}. \tag{3.10}$$

The formula $F(\vec{t}) = f(d, k)$ now follows from a straightforward computation. □

3.4 The second moment

This section gives an estimate on the expected number of proper colorings at a given Hamming distance from the planted coloring. This computation yields (3.2) of Lemma 3.0.6 as a corollary. It also reduces the proof of (3.3) to obtaining an estimate on the typical number of proper colorings near the planted coloring.

Before stating the main result, it seems worthwhile to review notation. Fix $n > 0$ with $n \in 2\mathbb{Z} \cap k\mathbb{Z}$. Fix an equitable 2-coloring $\chi : [n] \to \{0, 1\}$. This is the **planted coloring**. The planted model $\mathbb{P}_n^p$ is the uniform probability measure on the set $\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ of all uniform homomorphisms σ such that χ is σ-proper. Also let $Z_e : \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n)) \to \mathbb{N}$ be the number of equitable proper 2-colorings. For $\delta \in [0, 1]$, let $Z_\chi(\delta; \cdot) : \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n)) \to \mathbb{N}$ be the number of equitable proper 2-colorings $\widetilde{\chi}$ such that $|d_n(\chi, \widetilde{\chi}) - \delta| < 1/2n$ where d_n is the normalized Hamming distance defined by

$$d_n(\chi, \widetilde{\chi}) = n^{-1} \#\{v \in [n] : \ \chi(v) \neq \widetilde{\chi}(v)\}.$$

We will also write $Z_\chi(\delta; \sigma) = Z_\chi(\delta) = Z(\delta)$ when χ and/or σ are understood.

The main result of this section is:

Theorem 3.4.1. *With notation as above, for any $0 \leq \delta \leq 1$ such that $\delta n/2$ is an integer,*

$$\frac{1}{n} \log \mathbb{E}_n^p[Z(\delta)] = \psi_0(\delta) + O_\delta(n^{-1}) + O(n^{-1} \log(n))$$

(for $n \geq 2$) where

$$\psi_0(\delta) = (1-d)H(\delta, 1-\delta) + dH_0(\delta, 1-\delta) + \frac{d}{k}\log\left(1 - \frac{1-\delta_0^k - (1-\delta_0)^k}{2^{k-1}-1}\right),$$

δ_0 is defined to be the unique solution to

$$\delta_0 \frac{1 - 2^{2-k} + (\delta_0/2)^{k-1}}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} = \delta$$

and

$$\begin{aligned} H(\delta, 1-\delta) &:= -\delta\log\delta - (1-\delta)\log(1-\delta), \\ H_0(\delta, 1-\delta) &:= -\delta\log\delta_0 - (1-\delta)\log(1-\delta_0). \end{aligned}$$

Moreover, the constant implicit in the error term $O(n^{-1}\log(n))$ may depend on k but not on δ. The constant implicit in the $O_\delta(n^{-1})$ term depends continuously on δ for $\delta \in (0, 1/2]$.

Remark 3.4.2. *If $\delta_0 = \delta$ then $\delta = 1/2$. In the general case, $\delta_0 = \delta + O(2^{-k})$. Theorem 3.4.1 parallels similar results in [27, 28] for the random hyper-graph $H_k(n, m)$. This is explained in more detail in the next subsection.*

The strategy behind the proof of Theorem 3.4.1 is as follows. We need to estimate the expected number of equitable colorings at distance δ from the planted coloring. By symmetry, it suffices to fix another coloring $\widetilde{\chi}$ that is at distance δ from the planted coloring and count the number of uniform homomorphisms σ such that both χ and $\widetilde{\chi}$ are proper with respect to G_σ. This can be handled one generator at a time. Moreover, only the orbit-partition induced by a generator is used in this computation. So, for fixed $\chi, \widetilde{\chi}$, we need to estimate the number of k-partitions of $[n]$ that are bi-chromatic under both χ and $\widetilde{\chi}$. To make this strategy precise, we need the next definitions.

Definition 3.4.3. *Let $\widetilde{\chi}$ be an equitable 2-coloring of $[n]$. An edge $P \subset [n]$ is* $(\chi, \widetilde{\chi})$**-bichromatic** *if $\chi(P) = \widetilde{\chi}(P) = \{0, 1\}$. Recall that a k-partition is a partition $\pi = \{P_1, \ldots, P_{n/k}\}$ of $[n]$ such that every part $P \in \pi$ has cardinality k. A k-partition π is* $(\chi, \widetilde{\chi})$**-bichromatic** *if every part $P \in \pi$ is $(\chi, \widetilde{\chi})$-bichromatic.*

Given a $(\chi, \widetilde{\chi})$-bichromatic edge $P \subset [n]$ of size k, there is a 2×2 matrix $\vec{e}(\widetilde{\chi}, P)$ defined by

$$\vec{e}_{i,j}(\widetilde{\chi}, P) = |P \cap \chi^{-1}(i) \cap \widetilde{\chi}^{-1}(j)|.$$

Let $\mathcal{E}$ denote the set of all such matrices (over all $P, \widetilde{\chi}$). This is a finite set. To be precise, $\mathcal{E}$ is the set of all 2×2 matrices $\vec{e} = (e_{ij})_{i,j=0,1}$ such that

- $e_{ij} \in \{0, 1, \ldots, k\}$ *for all* i, j
- $0 < e_{10} + e_{11} < k$
- $0 < e_{01} + e_{11} < k$
- $\sum_{i,j} e_{ij} = k$.

If π is a $(\chi, \widetilde{\chi})$-bichromatic k-partition then it induces a function $t_{\widetilde{\chi},\pi} : \mathcal{E} \to [0, 1]$ by

$$t_{\widetilde{\chi},\pi}(\vec{e}) = n^{-1} \# \left\{ P \in \pi : \ \vec{e} = \vec{e}(\widetilde{\chi}, P) \right\}.$$

Let $\mathcal{T}$ be the set of all functions $t : \mathcal{E} \to [0, 1]$ satisfying

- $\sum_{\vec{e} \in \mathcal{E}} t(\vec{e}) = 1/k$,
- $\sum_{\vec{e} \in \mathcal{E}} (e_{10} + e_{11}) t(\vec{e}) = 1/2$,
- $\sum_{\vec{e} \in \mathcal{E}} (e_{01} + e_{11}) t(\vec{e}) = 1/2$.

Also let $\mathcal{T}_n$ be the set of $t \in \mathcal{T}$ such that $nt(\vec{e})$ is integer-valued for each $\vec{e} \in \mathcal{E}$. A k-partition π has **type** $(\chi, \widetilde{\chi}, t)$ *if $t = t_{\widetilde{\chi},\pi}$.*

Lemma 3.4.4. *Given an equitable 2-coloring $\widetilde{\chi} : [n] \to \{0,1\}$, let $p^{\widetilde{\chi}} = (p_{ij}^{\widetilde{\chi}})$ be the 2×2 matrix*

$$p_{ij}^{\widetilde{\chi}} = (1/n)|\chi^{-1}(i) \cap \widetilde{\chi}^{-1}(j)|.$$

Then

$$p^{\widetilde{\chi}} = \begin{bmatrix} 1/2 - d_n(\chi, \widetilde{\chi})/2 & d_n(\chi, \widetilde{\chi})/2 \\ d_n(\chi, \widetilde{\chi})/2 & 1/2 - d_n(\chi, \widetilde{\chi})/2 \end{bmatrix}$$

In particular, $p^{\widetilde{\chi}}$ is determined by the Hamming distance $d_n(\chi, \widetilde{\chi})$.

Proof. Let $p = p^{\widetilde{\chi}}$. The lemma follows from this system of linear equations:

$$\begin{aligned} 1/2 &= p_{01} + p_{11} \\ 1/2 &= p_{10} + p_{11} \\ d_n(\chi, \widetilde{\chi}) &= p_{01} + p_{10} \\ 1 &= p_{00} + p_{01} + p_{10} + p_{11}. \end{aligned}$$

The first two occur because both χ and $\widetilde{\chi}$ are equitable. The third follows from the definition of normalized Hamming distance and the last holds because $\{\chi^{-1}(i) \cap \widetilde{\chi}^{-1}(j)\}_{i,j \in \{0,1\}}$ partitions $[n]$. □

For $t \in \mathcal{T}$, define the 2×2 matrix $p^t = (p_{ij}^t)$ by

$$p_{ij}^t := \sum_{\vec{e} \in \mathcal{E}} e_{ij} t(\vec{e}).$$

If π is a k-partition that has type $(\chi, \widetilde{\chi}, t)$ (for some equitable $\widetilde{\chi}$) then $p^{\widetilde{\chi}} = p^t$. This motivates the definition.

The main combinatorial estimate we will need is:

Lemma 3.4.5. *Let $t \in \mathcal{T}_n$ and $\widetilde{\chi} : [n] \to \{0,1\}$ be equitable. Suppose $p^t = p^{\widetilde{\chi}}$. Let*

$g(\widetilde{\chi}, t)$ be the number of k-partitions of type $(\chi, \widetilde{\chi}, t)$. Also let

$$G(t) := (1 - 1/k)(\log(n) - 1) - H(p^t) - (1/k)\log(k!) + H(t) + \sum_{\vec{e}} t(\vec{e}) \log \binom{k}{\vec{e}}$$

where $\binom{k}{\vec{e}}$ is the multinomial $\frac{k!}{e_{00}!e_{01}!e_{10}!e_{11}!}$. Then

$$(1/n)\log g(\widetilde{\chi}, t) = G(t) + O(n^{-1}\log(n))$$

(for $n \geq 2$) where the constant implicit in the error term depends on k but not on $\widetilde{\chi}$ or t.

Proof. The following algorithm constructs all such partitions with no duplications:

Step 1. Choose a partition $\{Q_{ij}^{\vec{e}} : \ i, j \in \{0,1\}, \vec{e} \in \mathcal{E}\}$ of $\chi^{-1}(i) \cap \widetilde{\chi}^{-1}(j)$ such that

$$|Q_{ij}^{\vec{e}}| = e_{ij} t(\vec{e}) n.$$

Step 2. For $i, j \in \{0,1\}$ and $\vec{e} \in \mathcal{E}$, choose an unordered partition $\pi_{ij}^{\vec{e}}$ of $Q_{ij}^{\vec{e}}$ into $t(\vec{e})n$ sets of size e_{ij}.

Step 3. For $i, j \in \{0,1\}$ with $(i,j) \neq (0,0)$ and $\vec{e} \in \mathcal{E}$, choose a bijection $\alpha_{ij}^{\vec{e}} : \pi_{00}^{\vec{e}} \to \pi_{ij}^{\vec{e}}$.

Step 4. The k-partition consists of all sets of the form $P \cup \bigcup_{i,j \in \{0,1\}, (i,j) \neq (0,0)} \alpha_{ij}^{\vec{e}}(P)$ over all $P \in \pi_{00}^{\vec{e}}$ and $\vec{e} \in \mathcal{E}$.

The number of choices in Step 1 is

$$\prod_{i,j \in \{0,1\}} |\chi^{-1}(i) \cap \widetilde{\chi}^{-1}(j)|! \prod_{\vec{e} \in \mathcal{E}} (e_{ij} t(\vec{e}) n)!^{-1}.$$

The combined number of choices in Steps 1 and 2 is

$$\prod_{i,j\in\{0,1\}} |\chi^{-1}(i)\cap\widetilde{\chi}^{-1}(j)|!\prod_{\vec{e}\in\mathcal{E}} e_{ij}!^{-t(\vec{e})n}(t(\vec{e})n)!^{-1}$$
$$=\left(\prod_{\vec{e}\in\mathcal{E}}(t(\vec{e})n)!\right)^{-4}\prod_{i,j\in\{0,1\}} |\chi^{-1}(i)\cap\widetilde{\chi}^{-1}(j)|!\prod_{\vec{e}\in\mathcal{E}} e_{ij}!^{-t(\vec{e})n}.$$

The number of choices in Step 3 is $\prod_{\vec{e}\in\mathcal{E}}(t(\vec{e})n)!^3$. So

$$g(\widetilde{\chi},t)=\left(\prod_{i,j\in\{0,1\}} |\chi^{-1}(i)\cap\widetilde{\chi}^{-1}(j)|!\right)\left(\prod_{\vec{e}\in\mathcal{E}}(t(\vec{e})n)!\right)^{-1}\left(\prod_{i,j\in\{0,1\}}\prod_{\vec{e}\in\mathcal{E}} e_{ij}!^{-t(\vec{e})n}\right).$$

An application of Stirling's formula gives

$$(1/n)\log g(\widetilde{\chi},t)=(1-1/k)(\log(n)-1)-H(p^t)+H(t)-\sum_{\vec{e},i,j} t(\vec{e})\log(e_{ij}!)+O(n^{-1}\log(n))$$

(for $n\geq 2$) where the constant implicit in the error term depends on k but not on $\widetilde{\chi}$ or t.

Since $\sum_{\vec{e}} t(\vec{e})=1/k$,

$$\sum_{\vec{e}} t(\vec{e})\log\binom{k}{\vec{e}}=(1/k)\log(k!)-\sum_{\vec{e},i,j} t(\vec{e})\log(e_{ij}!).$$

Substitute this into the formula above to finish the lemma.

□

Next we use Lagrange multipliers to maximize $G(t)$. To be precise, for $\delta\in[0,1]$, let $\mathcal{T}(\delta)$ be the set of all $t\in\mathcal{T}$ such that $p_{01}^t=\delta/2$. Define $\mathcal{T}_n(\delta)=\mathcal{T}(\delta)\cap\mathcal{T}_n$. To motivate this definition, observe that if $\widetilde{\chi}$ is an equitable 2-coloring and $\delta=d_n(\chi,\widetilde{\chi})$ then $p_{01}^{\widetilde{\chi}}=\delta/2$. So if π is a k-partition with type $(\chi,\widetilde{\chi},t)$ then $p_{01}^t=\delta/2$.

Lemma 3.4.6. *Let $\delta \in [0,1]$. Then there exists a unique $s_\delta \in \mathcal{T}(\delta)$ such that*

$$\max_{t \in \mathcal{T}(\delta)} G(t) = G(s_\delta).$$

Moreover, if δ_0, $C > 0$ and $t_\delta \in \mathcal{T}(\delta)$ are defined by

$$\begin{aligned} \frac{\delta}{2} &= \frac{\delta_0}{2} \frac{1 - 2^{2-k} + (\delta_0/2)^{k-1}}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \\ C &= \frac{1}{k[1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k]} \\ t_\delta(\vec{e}) &= C \left(\frac{1-\delta_0}{2}\right)^{e_{00}+e_{11}} \left(\frac{\delta_0}{2}\right)^{e_{01}+e_{10}} \binom{k}{\vec{e}} \end{aligned}$$

then $s_\delta = t_\delta$.

Proof. Define $F : \mathcal{T} \to \mathbb{R}$ by

$$F(t) = H(t) + \sum_{\vec{e}} t(\vec{e}) \log \binom{k}{\vec{e}}.$$

For all $t \in \mathcal{T}(\delta)$, $G(t) - F(t)$ is constant in t. Therefore, it suffices to prove the lemma with F in place of G.

The function F is concave over $t \in \mathcal{T}(\delta)$. This implies the existence of a unique $s_\delta \in \mathcal{T}(\delta)$ such that

$$\max_{t \in \mathcal{T}(\delta)} F(t) = F(s_\delta).$$

By definition, $\mathcal{T}(\delta)$ is the set of all functions $t : \mathcal{E} \to [0,1]$ satisfying

$$\begin{aligned} 1/k &= \sum_{\vec{e} \in \mathcal{E}} t(\vec{e}) \\ p_{ij} &= \sum_{\vec{e} \in \mathcal{E}} e_{ij} t(\vec{e}), \end{aligned}$$

where $p = (p_{ij})$ is the matrix

$$p = \begin{bmatrix} 1/2 - \delta/2 & \delta/2 \\ \delta/2 & 1/2 - \delta/2 \end{bmatrix}.$$

For any $\vec{e} \in \mathcal{E}$,

$$\frac{\partial F}{\partial t(\vec{e})} = -\log t(\vec{e}) - 1 + \log \binom{k}{\vec{e}}. \tag{3.11}$$

Since this is positive infinity when $t(\vec{e}) = 0$, s_δ must lie in the interior of $\mathcal{T}(\delta)$. By the method of Lagrange multipliers there exists $\lambda \in \mathbb{R}$ and a 2×2 matrix $\vec{\mu}$ such that

$$\frac{\partial F}{\partial t(\vec{e})}(s_\delta) = \lambda + \vec{\mu} \cdot \vec{e}. \tag{3.12}$$

Evaluate (3.11) at s_δ, use (3.12) and solve for $s_\delta(\vec{e})$ to obtain

$$s_\delta(\vec{e}) = C_0 \binom{k}{\vec{e}} x_{00}^{e_{00}} x_{01}^{e_{01}} x_{10}^{e_{10}} x_{11}^{e_{11}}$$

for some constants C_0, x_{ij}. In fact, since F is concave, s_δ is the unique critical point and so it is the only element of $\mathcal{T}(\delta)$ of this form. So it suffices to check that the purported t_δ given in the statement of the lemma has this form and that it is in $\mathcal{T}(\delta)$ as claimed. The former is immediate while the latter is a tedious but straightforward computation. For example, to check that $\sum_{\vec{e}} t_\delta(\vec{e}) = 1/k$, observe that, by the multinomial formula for any $(x_{ij})_{i,j \in \{0,1\}}$,

$$\begin{aligned} & \sum_{\vec{e} \in \mathcal{E}} \binom{k}{\vec{e}} x_{00}^{e_{00}} x_{01}^{e_{01}} x_{10}^{e_{10}} x_{11}^{e_{11}} \\ = \; & \Big[(x_{00} + x_{01} + x_{10} + x_{11})^k \\ & -(x_{00} + x_{01})^k - (x_{00} + x_{10})^k - (x_{11} + x_{01})^k - (x_{11} + x_{10})^k + x_{00}^k + x_{01}^k + x_{10}^k + x_{11}^k \Big]. \end{aligned}$$

Substitute $x_{00} = x_{11} = \frac{1-\delta_0}{2}$ and $x_{01} = x_{10} = \delta_0/2$ to obtain

$$\sum_{\vec{e}\in\mathcal{E}} t_\delta(\vec{e}) \;\; = \;\; C\left[1 - 4(1/2)^k + 2\left(\frac{1-\delta_0}{2}\right)^k + 2\left(\frac{\delta_0}{2}\right)^k\right] = 1/k.$$

The rest of the verification that $t_\delta \in \mathcal{T}(\delta)$ is left to the reader. □

Proof of Theorem 3.4.1. Let $\mathcal{E}(\delta)$ be the set of all equitable 2-colorings $\widetilde{\chi} : [n] \to \{0,1\}$ such that $d_n(\widetilde{\chi}, \chi) = \delta$. Also let $F_{\widetilde{\chi}} \subset \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$ be the set of all σ such that $\widetilde{\chi}$ is a proper 2-coloring of the hyper-graph G_σ. By linearity of expectation,

$$\mathbb{E}_n^p[Z_\chi(\delta)] = \sum_{\widetilde{\chi}\in\mathcal{E}(\delta)} \mathbb{P}_n^u(F_{\widetilde{\chi}}|F_\chi).$$

The cardinality of $\mathcal{E}(\delta)$ is $\binom{n/2}{\delta n/2}^2$. By Stirling's formula

$$n^{-1}\log\binom{n/2}{\delta n/2}^2 = H(\delta, 1-\delta) + O(n^{-1}\log(n)). \tag{3.13}$$

We have $\mathbb{P}_n^u(F_{\widetilde{\chi}}|F_\chi)$ is the same for all $\widetilde{\chi} \in \mathcal{E}(\delta)$. This follows by noting that the distribution of hyper-graphs in the planted model is invariant under any permutation which fixes χ. If η, η' are two configurations with $d_n(\eta, \chi) = d_n(\eta', \chi) = \delta$ then there is a permutation $\pi \in \mathrm{Sym}(n)$ which fixes χ and such that $\eta \circ \pi = \eta'$. To see this note that we simply need to find a $\pi \in \mathrm{Sym}(n)$ which maps the sets $\chi^{-1}(i) \cap \eta^{-1}(j)$ to $\chi^{-1}(i) \cap \eta'^{-1}(j)$ for each $i, j \in \{0,1\}$. Such a map exists since for each i, j the two sets have the same size. It follows that

$$n^{-1}\log\mathbb{E}_n^p[Z_\chi(\delta)] = H(\delta, 1-\delta) + n^{-1}\log\mathbb{P}_n^u(F_{\widetilde{\chi}}|F_\chi) + O(n^{-1}\log(n)) \tag{3.14}$$

for any fixed $\widetilde{\chi} \in \mathcal{E}(\delta)$.

For $1 \le i \le d$, let $F_{\chi,i}$ be the set of uniform homomorphisms σ such that the orbit-partition of $\sigma(s_i)$ is χ-bichromatic in the sense that $\chi(P) = \{0,1\}$ for every P

in the orbit-partition of $\sigma(s_i)$. Then the events $\{F_{\chi,i} \cap F_{\widetilde{\chi},i}\}_{i=1}^{d}$ are i.i.d. and

$$F_\chi \cap F_{\widetilde{\chi}} = \bigcap_{i=1}^{d} F_{\chi,i} \cap F_{\widetilde{\chi},i}.$$

Therefore,

$$\mathbb{P}_n^u(F_{\widetilde{\chi}}|F_\chi) = \frac{\mathbb{P}_n^u(F_{\widetilde{\chi},1} \cap F_{\chi,1})^d}{\mathbb{P}_n^u(F_\chi)}. \tag{3.15}$$

Note $\mathbb{P}_n^u(F_{\chi,1} \cap F_{\widetilde{\chi},1})$ is, up to sub-exponential factors, equal to the maximum of $g(\widetilde{\chi}, t)$ over $t \in \mathcal{T}_n(\delta)$ divided by the number of k-partitions of $[n]$. So equation (3.6) implies

$$\frac{1}{n}\log \mathbb{P}_n^u(F_{\widetilde{\chi},1} \cap F_{\chi,1}) = \max_{t \in \mathcal{T}_n(\delta)} \frac{1}{n}\log g(\widetilde{\chi}, t) - (1-1/k)(\log(n)-1) + (1/k)\log((k-1)!) + O(n^{-1}\log n).$$

So Lemma 3.4.5 implies

$$\frac{1}{n}\log \mathbb{P}_n^u(F_{\widetilde{\chi},1} \cap F_{\chi,1}) = \max_{t \in \mathcal{T}_n(\delta)} G(t) - (1-1/k)(\log(n)-1) + (1/k)\log((k-1)!) + O(n^{-1}\log n).$$

We apply Lemma 3.3.9 to $\mathcal{T}(\delta)$ to obtain the existence of $s_\delta^{(n)} \in \mathcal{T}_n(\delta)$ with $\|s_\delta^{(n)} - t_\delta\|_\infty < O_\delta(n^{-1})$ where the constant implicit in the $O_\delta(n^{-1})$ term depends continuously on δ for $\delta \in (0, 1/2]$. Since G is differentiable in a neighborhood of t_δ, we have

$$\max_{t \in \mathcal{T}_n(\delta)} G(t) = G(t_\delta) + O_\delta(n^{-1}).$$

So Lemma 3.4.6 implies

$$\begin{aligned} \frac{1}{n}\log \mathbb{P}_n^u(F_{\widetilde{\chi},1} \cap F_{\chi,1}) &= -H(\vec{p}) + H(t_\delta) - \sum_{i,j,\vec{e}} t_\delta(\vec{e})\log(e_{ij}!) + (1/k)\log(k-1)! \\ &\quad + O_\delta(n^{-1}) + O(n^{-1}\log(n)). \end{aligned}$$

Since $\vec{p} = (\delta/2, \delta/2, (1-\delta)/2, (1-\delta)/2)$, $H(\vec{p}) = H(\delta, 1-\delta) + \log(2)$. So

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}_n^u(F_{\widetilde{\chi},1} \cap F_{\chi,1}) &= -H(\delta,1-\delta) - \log(2) + H(t_\delta) - \sum_{i,j,\vec{e}} t_\delta(\vec{e}) \log(e_{ij}!) \\ &\quad + (1/k)\log(k-1)! + O_\delta(n^{-1}) + O(n^{-1}\log(n)). \quad (3.16)\end{aligned}$$

On the other hand, Theorem 3.3.2 implies

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}_n^u(F_\chi) &= \frac{1}{n}\log\left(\binom{n}{n/2}^{-1} \mathbb{E}_n^u[Z_e(\sigma)]\right) \\ &= (d/k)\log(1-2^{1-k}) + O(n^{-1}\log(n)).\end{aligned}$$

Combine this result with (3.14), (3.15) and (3.16) to obtain

$$\begin{aligned}n^{-1}\log \mathbb{E}_n^p[Z_\chi(\delta)] &= (1-d)H(\delta,1-\delta) - d\log(2) + dH(t_\delta) - d\sum_{i,j,\vec{e}} t_\delta(\vec{e})\log(e_{ij}!) \\ &\quad + (d/k)\log(k-1)! - (d/k)\log(1-2^{1-k}) + O_\delta(n^{-1}) + O(n^{-1}\log(n)).\end{aligned}$$

Since $\sum_{\vec{e}} t_\delta(\vec{e}) = 1/k$,

$$\sum_{\vec{e}} t_\delta(\vec{e}) \log\binom{k}{\vec{e}} = (1/k)\log k! - \sum_{i,j,\vec{e}} t_\delta(\vec{e})\log(e_{ij}!).$$

Substitute this into the previous equation to obtain

$$n^{-1}\log \mathbb{E}_n^p[Z_\chi(\delta)] = \psi_0(\delta) + O_\delta(n^{-1}) + O(n^{-1}\log(n))$$

where

$$\begin{aligned}\psi_0(\delta) &= (1-d)H(\delta,1-\delta) - d\log(2) + dH(t_\delta) + d\sum_{\vec{e}\in\mathcal{E}} t_\delta(\vec{e})\log\binom{k}{\vec{e}} \\ &\quad -(d/k)\log k - (d/k)\log(1-2^{1-k}).\end{aligned}$$

Observe that in every estimate above, the constant implicit in the error term does not depend on δ. To finish the lemma, we need only simplify the expression for ψ_0.

By Lemma 3.4.6,

$$\begin{aligned} H(t_\delta) &= -\sum_{\vec{e}} t_\delta(\vec{e}) \log t_\delta(\vec{e}) \\ &= -\sum_{\vec{e}} t_\delta(\vec{e}) \left(\log C + (e_{00}+e_{11}) \log\left(\frac{1-\delta_0}{2}\right) + (e_{01}+e_{10}) \log\left(\frac{\delta_0}{2}\right) + \log \binom{k}{\vec{e}} \right) \\ &= -(1/k)(\log C) - (1-\delta)\log(1-\delta_0) - \delta \log(\delta_0) + \log 2 - \sum_{\vec{e}} t_\delta(\vec{e}) \log \binom{k}{\vec{e}} \\ &= -(1/k)(\log C) + H_0(\delta, 1-\delta) + \log 2 - \sum_{\vec{e}} t_\delta(\vec{e}) \log \binom{k}{\vec{e}}. \end{aligned}$$

Combined with the previous formula for ψ_0, this implies

$$\psi_0(\delta) = (1-d)H(\delta, 1-\delta) - (d/k)\log C + dH_0(\delta, 1-\delta) - (d/k)\log k - (d/k)\log(1-2^{1-k}).$$

To simplify further, use the formula for C in Lemma 3.4.6 to obtain

$$\begin{aligned} -(d/k)\left(\log C + \log k + \log(1-2^{1-k})\right) &= (d/k)\log \frac{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k}{1-2^{1-k}} \\ &= (d/k)\log\left(1 - \frac{1-\delta_0^k - (1-\delta_0)^k}{2^{k-1}-1}\right). \end{aligned}$$

Thus $\psi_0(\delta) = (1-d)H(\delta, 1-\delta) + dH_0(\delta, 1-\delta) + \frac{d}{k}\log(1 - \frac{1-\delta_0^k-(1-\delta_0)^k}{2^{k-1}-1})$.

□

3.4.1 Analysis of ψ_0 and the proof of Lemma 3.0.6 inequality (3.2)

Theorem 3.4.1 reduces inequality (3.2) to analyzing the function ψ_0. A related function ψ, defined by

$$\psi(x) := H(x, 1-x) + \frac{d}{k} \log\left(1 - \frac{1 - x^k - (1-x)^k}{2^{k-1} - 1}\right)$$

has been analyzed in [27, 28]. It is shown there $\psi(x)$ is the exponential rate of growth of the number of proper colorings at normalized distance x from the planted coloring in the model $H_k(n, m)$. Moreover, if $r = d/k$ is close to $\frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2$ then the global maximum of $\psi(x)$ is attained at some $x \in (0, 2^{-k/2})$. Moreover, ψ has a local maximum at $x = 1/2$ and is symmetric around $x = 1/2$. It is negative in the region $(2^{-k/2}, 1/2 - 2^{-k/2})$. We will not need these facts directly, and mention them only for context, especially because we will obtain similar results for ψ_0.

The relevance of ψ to ψ_0 lies in the fact that

$$\psi_0(\delta) = \psi(\delta_0) - (H(\delta_0, 1-\delta_0) - H_0(\delta, 1-\delta)) + (d-1)\left[H_0(\delta, 1-\delta) - H(\delta, 1-\delta)\right] \quad (3.17)$$

As an aside, note that $H_0(\delta, 1-\delta) - H(\delta, 1-\delta)$ is the Kullback-Leibler divergence of the distribution $(\delta, 1-\delta)$ with respect to $(\delta_0, 1-\delta_0)$.

To prove inequality (3.2), we first estimate the difference $\psi_0(\delta) - \psi(\delta_0)$ and then estimate $\psi(\delta_0)$. Because the estimates we obtain are useful in the next subsection, we prove more than what is required for just inequality (3.2).

Lemma 3.4.7. *Suppose* $0 \leq \delta_0 \leq 1/2$. *Define* $\varepsilon \geq 0$ *by* $\delta = \delta_0(1-\varepsilon)$. *Then*

$$\begin{aligned} H(\delta_0, 1-\delta_0) - H_0(\delta, 1-\delta) &= \delta_0 \varepsilon \log\left(\frac{1-\delta_0}{\delta_0}\right) \geq 0, \\ H_0(\delta, 1-\delta) - H(\delta, 1-\delta) &= O(\delta_0 \varepsilon^2), \\ \varepsilon &= O(2^{-k}), \\ (1-\delta)_0 &= 1-\delta_0. \end{aligned}$$

The last equation implies $\psi_0(1-\delta) = \psi_0(\delta)$.

Proof. The first equality follows from:

$$\begin{aligned} H(\delta_0, 1-\delta_0) - H_0(\delta, 1-\delta) &= -\delta_0 \log \delta_0 - (1-\delta_0)\log(1-\delta_0) + \delta \log \delta_0 + (1-\delta)\log(1-\delta_0) \\ &= (\delta_0 - \delta)\log(1/\delta_0) + (\delta_0 - \delta)\log(1-\delta_0) \\ &= \delta_0 \varepsilon \log\left(\frac{1-\delta_0}{\delta_0}\right). \end{aligned}$$

The second estimate follows from:

$$\begin{aligned} H_0(\delta, 1-\delta) - H(\delta, 1-\delta) &= \delta\left(\log \delta - \log \delta_0\right) + (1-\delta)\left(\log(1-\delta) - \log(1-\delta_0)\right) \\ &= \delta \log(1-\varepsilon) + (1-\delta)\log\left(\frac{1-\delta}{1-\delta_0}\right) \\ &= -\delta\varepsilon + (1-\delta)\log\left(1 + \frac{\delta_0 \varepsilon}{1-\delta_0}\right) + O(\delta_0 \varepsilon^2) \\ &= -\delta\varepsilon + \delta_0 \varepsilon + O(\delta_0 \varepsilon^2) = O(\delta_0 \varepsilon^2). \end{aligned}$$

The third estimate follows from:

$$\begin{aligned}
\varepsilon &= 1 - \frac{\delta}{\delta_0} \\
&= 1 - \frac{1 - 2^{2-k} + (\delta_0/2)^{k-1}}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \\
&= \frac{2(\delta_0/2)^k - (\delta_0/2)^{k-1} + 2((1-\delta_0)/2)^k}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \\
&= \frac{(\delta_0/2)^{k-1}(\delta_0 - 1) + 2((1-\delta_0)/2)^k}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \\
&= 2^{1-k} \cdot (1-\delta_0) \cdot \frac{(1-\delta_0)^{k-1} - \delta_0^{k-1}}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k}. \qquad (3.18)
\end{aligned}$$

The denominator is $1 + O(2^{-k})$ and the numerator is $O(2^{-k})$. The result follows.

The last equation follows from:

$$\begin{aligned}
1 - \delta &= 1 - \delta_0 \left(\frac{1 - 2^{2-k} + (\delta_0/2)^{k-1}}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \right) \\
&= \frac{1 - 2^{2-k} + 2(\delta_0/2)^k + 2\left((1-\delta_0)/2)^k - \delta_0(1 - 2^{2-k} + (\delta_0/2)^{k-1}\right)}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \\
&= \frac{(1-\delta_0)\left(1 - 2^{2-k} + ((1-\delta_0)/2)^{k-1}\right)}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k}.
\end{aligned}$$

The last expression shows that $(1-\delta)_0 = 1 - \delta_0$.

□

Lemma 3.4.8. *Let $0 \leq \eta$ be $O_k(1)$. If*

$$r = d/k = \frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2 + \eta$$

then

$$\begin{aligned}
f(d,k) &= \psi(1/2) = \psi_0(1/2) = (1 - 2\eta)2^{-k} + O(2^{-2k}) \\
\psi(2^{-k}) &= 2^{-k} + O(k^2 2^{-2k}).
\end{aligned}$$

In particular, if k is sufficiently large then $\psi(2^{-k}) > f(d,k)$.

Proof. By direct inspection $f(d,k) = \psi(1/2) = \psi_0(1/2)$. By Taylor series expansion, $\log(1-2^{1-k}) = -2^{1-k} - 2^{1-2k} + O(2^{-3k})$. So

$$\begin{aligned} f(d,k) &= \log(2) + r\log(1-2^{1-k}) \\ &= \log(2) + \left(\frac{\log(2)}{2}\cdot 2^k - (1+\log(2))/2 + \eta\right)\left(-2^{1-k} - 2^{1-2k}\right) + O(r2^{-3k}) \\ &= (1-2\eta)2^{-k} + O(2^{-2k}). \end{aligned}$$

Next we estimate $\psi(2^{-k})$. For convenience, let $x = 2^{-k}$. Then

$$1 - x^k - (1-x)^k = k\cdot 2^{-k} + O(k^2 2^{-2k}).$$

Since $\log(1-x) = -x - x^2/2 + O(x^3)$,

$$\log\left(1 - \frac{1-x^k-(1-x)^k}{2^{k-1}-1}\right) = -2k\cdot 2^{-2k} + O(k^2 2^{-3k}).$$

So

$$r\log\left(1 - \frac{1-x^k-(1-x)^k}{2^{k-1}-1}\right) = -k\log(2)\cdot 2^{-k} + O(k^2 2^{-2k}).$$

Also,

$$H(x,1-x) = (k\log(2)+1)\cdot 2^{-k} + O(2^{-2k}).$$

Add these together to obtain

$$\psi(2^{-k}) = 2^{-k} + O(k^2 2^{-2k}).$$

□

Corollary 3.4.9. *Inequality (3.2) of Lemma 3.0.6 is true. To be precise, let* $0 < \eta_0$

be constant with respect to k. Then for all sufficiently large k (depending on η_0), if

$$r = d/k = \frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2 + \eta$$

for some $\eta \geq \eta_0$ with $\eta = O_k(1)$ then

$$f(d,k) < \liminf_{n\to\infty} n^{-1} \log \mathbb{E}_n^p[Z(\sigma)].$$

Proof. By definition,

$$n^{-1} \log \mathbb{E}_n^p[Z(\sigma)] \geq \max_{\delta\in[0,1/2]} n^{-1} \log \mathbb{E}_n^p[Z(\delta)].$$

By Theorem 3.4.1,

$$n^{-1} \log \mathbb{E}_n^p[Z(\sigma)] \geq \max_{\delta\in C_n} \psi_0(\delta) + O_\delta(n^{-1}) + O(n^{-1}\log(n))$$

where C_n is the set of $\delta \in [0,1/2]$ such that $\delta n/2$ is an integer. Because $\psi_0 : [0,1/2] \to \mathbb{R}$ is continuous, $\lim_{n\to\infty} \max_{\delta\in C_n} \psi_0(\delta) = \max_{\delta\in[0,1/2]} \psi_0(\delta)$. The constant implicit in the $O_\delta(n^{-1})$ depends continuously on δ for $\delta \in (0,1/2]$, while the constant implicit in the $O(n^{-1}\log(n))$ error term does not depend on δ. So

$$\liminf_{n\to\infty} n^{-1} \log \mathbb{E}_n^p[Z(\sigma)] \geq \max_{\delta\in[\upsilon,1/2]} \psi_0(\delta)$$

for any $0 < \upsilon < 1/2$. Because ψ_0 is continuous,

$$\liminf_{n\to\infty} n^{-1} \log \mathbb{E}_n^p[Z(\sigma)] \geq \max_{\delta\in[0,1/2]} \psi_0(\delta). \tag{3.19}$$

Because $H_0(\delta, 1-\delta) - H(\delta, 1-\delta) \geq 0$ (since it is a Kullback-Liebler divergence), the

first equality of Lemma 3.4.7 implies

$$\begin{aligned}\psi_0(\delta) &= \psi(\delta_0) - (H(\delta_0, 1-\delta_0) - H_0(\delta, 1-\delta)) + (d-1)\left[H_0(\delta, 1-\delta) - H(\delta, 1-\delta)\right] \\ &\geq \psi(\delta_0) - \delta_0 \varepsilon \log\left(\frac{1-\delta_0}{\delta_0}\right).\end{aligned}$$

By Lemma 3.4.8, $\psi(2^{-k}) = f(d,k) + 2\eta 2^{-k} + O(k^2 2^{-2k})$. By Lemma 3.4.7, $\varepsilon = O(2^{-k})$. As δ varies over $[0, 1/2]$, δ_0 also varies over $[0, 1/2]$, so there exists δ such that $\delta_0 = 2^{-k}$. For this value of δ,

$$\psi_0(\delta) \geq \psi(2^{-k}) - 2^{-k}\varepsilon \log\left(\frac{1-2^{-k}}{2^{-k}}\right) \geq f(d,k) + 2\eta 2^{-k} + O(k^2 2^{-2k}).$$

Combined with (3.19) this implies the Corollary. □

In the next subsection, we will need the following result.

Proposition 3.4.10. *Let*

$$0 < \eta_0 < (1 - \log 2)/2.$$

Then there exists k_0 (depending on η_0) such that for all $k \geq k_0$ if

$$r = d/k = \frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2 + \eta$$

for some $\eta \in [\eta_0, (1 - \log 2)/2)$ then in the interval $[2^{-k/2}, 1 - 2^{-k/2}]$, ψ_0 attains its unique maximum at 1/2. That is,

$$\max\{\psi_0(\delta) : \ 2^{-k/2} \leq \delta \leq 1 - 2^{-k/2}\} = \psi_0(1/2) = f(d,k) = \log(2) + r\log(1 - 2^{1-k})$$

and if $\delta \in [2^{-k/2}, 1 - 2^{-k/2}]$ and $\delta \neq 1/2$ then $\psi_0(\delta) < \psi_0(1/2)$.

Proof. By Lemma 3.4.7, it suffices to restrict δ to the interval $[2^{-k/2}, 1/2]$ (because $\psi_0(\delta) = \psi_0(1-\delta)$). So we will assume $\delta \in [2^{-k/2}, 1/2]$ without further mention.

Define ψ_1 by

$$\psi_1(\delta_0) = \frac{d}{k} \log\left(1 - \frac{1 - \delta_0^k - (1-\delta_0)^k}{2^{k-1} - 1}\right).$$

Observe

$$\begin{aligned} \psi_1(\delta_0) &= r\left(-\frac{1 - \delta_0^k - (1-\delta_0)^k}{2^{k-1} - 1} + O(4^{-k})\right) \\ &= -\log(2)[1 - (1-\delta_0)^k] + O(2^{-k}). \end{aligned}$$

By (3.17) and the first inequality of Lemma 3.4.7,

$$\begin{aligned} \psi_0(\delta) &\leq \psi(\delta_0) + (d-1)[H_0(\delta, 1-\delta) - H(\delta, 1-\delta)] \\ &= H(\delta_0, 1-\delta_0) + \psi_1(\delta_0) + (d-1)[H_0(\delta, 1-\delta) - H(\delta, 1-\delta)]. \quad (3.20) \end{aligned}$$

Moreover, $(d-1) = O(k2^k)$ and, by Lemma 3.4.7, $H_0(\delta, 1-\delta) - H(\delta, 1-\delta) = O(\delta_0 4^{-k})$. Therefore,

$$\begin{aligned} \psi_0(\delta) &\leq H(\delta_0, 1-\delta_0) + \psi_1(\delta_0) + O(\delta_0 k 2^{-k}) \\ &\leq H(\delta_0, 1-\delta_0) - \log(2)[1 - (1-\delta_0)^k] + O((\delta_0 k + 1)2^{-k}). \quad (3.21) \end{aligned}$$

Observe that $\delta_0 \geq \delta$. We divide the rest of the proof into five cases depending on where δ_0 lies in the interval $[2^{-k/2}, 1/2]$.

Case 1. Suppose $2^{-k/2} \leq \delta_0 \leq \frac{1}{2k}$. We claim that $\psi_0(\delta) < 0$. Note $-\log(\delta_0) \leq (k/2)\log(2)$ and $-(1-\delta_0)\log(1-\delta_0) = \delta_0 + O(\delta_0^2)$. So

$$\begin{aligned} H(\delta_0, 1-\delta_0) &= -\delta_0 \log \delta_0 - (1-\delta_0)\log(1-\delta_0) \\ &\leq \delta_0 (k/2)\log(2) + \delta_0 + O(\delta_0^2). \end{aligned}$$

By Taylor series expansion,

$$1-(1-\delta_0)^k \geq k\delta_0 - \binom{k}{2}\delta_0^2 \geq 3k\delta_0/4.$$

So by (3.21)

$$\begin{aligned}\psi_0(\delta) &\leq \delta_0(k/2)\log(2) + \delta_0 - 3k\delta_0\log(2)/4 + O(\delta_0^2)\\ &= \delta_0[1-k\log(2)/4] + O(\delta_0^2).\end{aligned}$$

Thus $\psi_0(\delta) < 0$ if k is sufficiently large.

Case 2. Let $0 < \xi_0 < 1/2$ be a constant such that $H(\xi_0, 1-\xi_0) < \log(2)(1-e^{-1/2})$. Suppose $\frac{1}{2k} \leq \delta_0 \leq \xi_0$. We claim that $\psi_0(\delta) < 0$ if k is sufficiently large.

By monotonicity, $H(\delta_0, 1-\delta_0) \leq H(\xi_0, 1-\xi_0)$. Since $1-x \leq e^{-x}$ (for $x > 0$),

$$[1-(1-\delta_0)^k] \geq 1 - e^{-k\delta_0} \geq 1 - e^{-1/2}.$$

By (3.21),

$$\psi_0(\delta) \leq H(\xi_0, 1-\xi_0) - \log(2)(1-e^{-1/2}) + O(k2^{-k}).$$

This implies the claim.

Case 3. Let ξ_1 be a constant such that $\max(\xi_0, 1/3) < \xi_1 < 1/2$. Suppose $\xi_0 \leq \delta_0 \leq \xi_1$. We claim that $\psi_0(\delta) < 0$ for all sufficiently large k (depending on ξ_1).

By (3.21),

$$\begin{aligned}\psi_0(\delta) &\leq H(\xi_1, 1-\xi_1) - \log(2)[1-(1-\delta_0)^k] + O(k2^{-k})\\ &\leq H(\xi_1, 1-\xi_1) - \log(2) + O((1-\xi_0)^k).\end{aligned}$$

This proves the claim.

Case 4. We claim that if $\xi_1 \leq \delta_0 \leq 0.5 - 2^{-k}$ then $\psi_0(\delta) < f(d,k)$ for all sufficiently large k (independent of the choice of ξ_1).

Recall that we define ε by $\delta = \delta_0(1-\varepsilon)$. By (3.18),

$$\begin{aligned} \varepsilon &= 2^{1-k} \cdot (1-\delta_0) \cdot \frac{(1-\delta_0)^{k-1} - \delta_0^{k-1}}{1 - 2^{2-k} + 2(\delta_0/2)^k + 2((1-\delta_0)/2)^k} \\ &\leq 2^{1-k}(1-\xi_1)^k + O\left(4^{-k}\right) \leq 2 \cdot 3^{-k} \end{aligned}$$

since $\xi_1 > 1/3$, assuming k is sufficiently large.

The assumption on r implies $d = O\left(k2^k\right)$. So the second equality of Lemma 3.4.7 implies

$$(d-1)[H_0(\delta, 1-\delta) - H(\delta, 1-\delta)] = O\left(k4.5^{-k}\right).$$

Through elementary but messy calculus computations one may show using the fact that $r = O(2^k)$ that

$$\begin{aligned} \psi'(1/2) &= 0 \\ \psi''(1/2) &= -4 + O(k^2 2^{-k}) \\ \psi'''(x) &= \frac{1}{x^2} - \frac{1}{(1-x)^2} + O(k^3(2/3)^k) \text{ for any } x \in [1/3, 2/3]. \end{aligned}$$

By Taylor's theorem,

$$\psi(\delta_0) = \psi(1/2) - (2 + o_k(1))(\delta_0 - 1/2)^2 + \frac{1}{6}\left(\frac{1}{\upsilon^2} - \frac{1}{(1-\upsilon)^2} + o_k(1)\right)(\delta_0 - 1/2)^3 \quad (3.22)$$

for some $\delta_0 \leq \upsilon \leq 1/2$. Since $\delta_0 \geq 1/3$, we have $1/2 - \delta_0 \leq 1/6$. Furthermore, since $x \mapsto \frac{1}{x^2} - \frac{1}{(1-x)^2}$ is monotone decreasing on $(0, 1/2]$, we have $\frac{1}{\upsilon^2} - \frac{1}{(1-\upsilon)^2} \leq \frac{1}{(1/3)^2} - \frac{1}{(2/3)^2} = \frac{27}{4}$. Thus

$$\frac{\frac{1}{6}\left(\frac{1}{\upsilon^2} - \frac{1}{(1-\upsilon)^2}\right)(1/2 - \delta_0)^3}{2(1/2-\delta_0)^2} \leq \frac{1}{12} \cdot \frac{27}{4} \cdot \frac{1}{6} < 1$$

and for sufficiently large k we have $\frac{1}{6}\left(\frac{1}{\upsilon^2}-\frac{1}{(1-\upsilon)^2}+o_k(1)\right)(1/2-\delta_0)^3<(2+o_k(1))(1/2-\delta_0)^2$. Since $\psi(1/2)=f(d,k)$, (3.20) implies

$$\psi_0(\delta)\leq f(d,k)-(2+o_k(1))(1/2-\delta_0)^2+\frac{1}{6}\left(\frac{1}{\upsilon^2}-\frac{1}{(1-\upsilon)^2}+o_k(1)\right)(1/2-\delta_0)^3+O\left(k4.5^{-k}\right)$$

is strictly less than $f(d,k)$ if k is sufficiently large.

Case 5. Suppose $0.5-2^{-k}\leq\delta_0<0.5$. Let $\gamma=0.5-\delta_0$. By (3.18),

$$\varepsilon=O\left(\left[(1/2+\gamma)^{k-1}-(1/2-\gamma)^{k-1}\right]2^{-k}\right).$$

Define $L(x):=(1/2+x)^{k-1}-(1/2-x)^{k-1}$. We claim that $L(\gamma)\leq\gamma$. Since $L(0)=0$, it suffices to show that $L'(x)\leq 1$ for all x with $|x|\leq 0.01$. An elementary calculation shows

$$L'(x)=(k-1)\left[(1/2+x)^{k-2}+(1/2-x)^{k-2}\right].$$

So $L'(x)\leq 1$ if $|x|\leq 0.01$ and k is sufficiently large. Altogether this proves $\varepsilon=O\left(\gamma 2^{-k}\right)$. So the second equality of Lemma 3.4.7 implies

$$(d-1)\left[H_0(\delta,1-\delta)-H(\delta,1-\delta)\right]=O\left(k2^{-k}\gamma^2\right).$$

By (3.22) and (3.20),

$$\psi_0(\delta)\leq f(d,k)-(2+o_k(1))\gamma^2+O\left(k2^{-k}\gamma^2\right).$$

This is strictly less than $f(d,k)$ if k is sufficiently large.

□

3.4.2 Reducing Lemma 3.0.6 inequality (3.3) to estimating the local cluster

As in the previous section, fix an equitable coloring $\chi : V_n \to \{0, 1\}$. Given a uniform homomorphism $\sigma \in \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$, the **cluster around** χ is the set

$$\mathcal{C}_\sigma(\chi) := \left\{\widetilde{\chi} \in Z_e(\sigma) : d_n(\chi, \widetilde{\chi}) \leq 2^{-k/2}\right\}.$$

We also call this the **local cluster** if χ is understood.

In §3.5 we prove:

Proposition 3.4.11. *Let $0 < \eta_0 < \eta_1 < (1 - \log 2)/2$. Then for all sufficiently large k (depending on η_0, η_1), if*

$$r := d/k = \frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2 + \eta$$

for some η with $\eta_0 \leq \eta \leq \eta_1$ then with high probability in the planted model, $|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)$. In symbols,

$$\lim_{n\to\infty} \mathbb{P}_n^\chi\big(|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)\big) = 1.$$

The rest of this section proves Lemma 3.0.6 inequality (3.3) from Proposition 3.4.11 and the second moment estimates from earlier in this section. So we assume the hypotheses of Proposition 3.4.11 without further mention.

We say that a coloring χ is **σ-good** if it is equitable and $|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e(\sigma))$. Let $S_g(\sigma)$ be the set of all σ-good proper colorings and let $Z_g(\sigma) = |S_g(\sigma)|$ be the number of σ-good proper colorings.

We will say a positive function $G(n)$ is **sub-exponential in** n if $\lim_{n\to\infty} n^{-1} \log G(n) = 0$. Also we say functions G and H are **asymptotic**, denoted by $G(n) \sim H(n)$, if $\lim_{n\to\infty} G(n)/H(n) = 1$. Similarly, $G(n) \lesssim H(n)$ if $\limsup_{n\to\infty} G(n)/H(n) \leq 1$.

Lemma 3.4.12. $\mathbb{E}_n^u(Z_g) \sim \mathbb{E}_n^u(Z_e) = F(n)\mathbb{E}_n^u(Z)$ *where* $F(n)$ *is sub-exponential in* n.

Proof. For brevity, let $\mathcal{H} = \mathrm{Hom}_{\mathrm{unif}}(\Gamma, \mathrm{Sym}(n))$. Let $\mathbb{P}_n^\chi$ be the probability operator in the planted model of χ. By definition,

$$\begin{aligned}
\mathbb{E}_n^u(Z_g) &= |\mathcal{H}|^{-1} \sum_{\sigma \in \mathcal{H}} Z_g(\sigma) = |\mathcal{H}|^{-1} \sum_{\sigma \in \mathcal{H}} \sum_{\chi: V \to \{0,1\}} 1_{S_g(\sigma)}(\chi) \\
&= \sum_{\chi} \mathbb{P}_n^u(\chi \in S_g(\sigma)) \\
&= \sum_{\chi \text{ equitable}} \mathbb{P}_n^u\left(|\mathcal{C}_\sigma(\chi)| \le \mathbb{E}_n^u(Z_e) | \chi \text{ proper}\right) \mathbb{P}_n^u(\chi \text{ proper}) \\
&= \sum_{\chi \text{ equitable}} \mathbb{P}_n^\chi\left(|\mathcal{C}_\sigma(\chi)| \le \mathbb{E}_n^u(Z_e)\right) \mathbb{P}_n^u(\chi \text{ proper}) \\
&\sim \sum_{\chi \text{ equitable}} \mathbb{P}_n^u\left(\chi \text{ proper}\right) = \mathbb{E}_n^u(Z_e)
\end{aligned}$$

where the asymptotic equality $\sim$ follows from Proposition 3.4.11. The equality $\mathbb{E}_n^u(Z_e) = F(n)\mathbb{E}_n^u(Z)$ holds by Theorem 3.3.2. □

Lemma 3.4.13. $\mathbb{E}_n^u(Z_g^2) \le C(n)\mathbb{E}_n^u(Z_g)^2$, *where* $C(n) = C(n,k,r)$ *is sub-exponential in* n.

Proof.

$$\begin{aligned}
\mathbb{E}_n^u(Z_g^2) &= |\mathcal{H}|^{-1} \sum_{\sigma \in \mathcal{H}} \left(\sum_{\chi} 1_{S_g(\sigma)}(\chi) \right)^2 \\
&= |\mathcal{H}|^{-1} \sum_{\sigma \in \mathcal{H}} \sum_{\chi, \widetilde{\chi}} 1_{S_g(\sigma)}(\chi) 1_{S_g(\sigma)}(\widetilde{\chi}) \\
&= \sum_{\chi, \widetilde{\chi}} \mathbb{P}_n^u\left(\chi \in S_g \text{ and } \widetilde{\chi} \in S_g\right) \\
&= \sum_{\chi, \widetilde{\chi}} \mathbb{P}_n^u(\chi \in S_g) \mathbb{P}_n^u(\widetilde{\chi} \in S_g | \chi \in S_g) \\
&= \sum_{\chi} \mathbb{P}_n^u(\chi \in S_g) \mathbb{E}_n^u(Z_g | \chi \in S_g). \qquad (3.23)
\end{aligned}$$

For a fixed $\chi \in S_g(\sigma)$ we analyze $\mathbb{E}_n^u(Z_g | \chi \in S_g)$ by breaking the colorings into those that are close (i.e. in the local cluster) and those that are far. So let $Z_g(\delta)$:

$\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n)) \to \mathbb{N}$ be the number of good proper colorings such that $d_n(\chi, \widetilde{\chi}) = \delta$. (We will also use $Z_e(\delta) = Z_\chi(\delta)$ for the analogous number of equitable proper colorings). Then

$$\mathbb{E}_n^u(Z_g|\chi \in S_g) \leq 2\mathbb{E}_n^u\left(\sum_{0\leq\delta\leq 2^{-k/2}} Z_g(\delta)\Big|\chi \in S_g\right) + 2\mathbb{E}_n^u\left(\sum_{2^{-k/2}<\delta\leq 1/2} Z_g(\delta)\Big|\chi \in S_g\right) \quad (3.24)$$

The coefficient 2 above accounts for the following symmetry: if $\widetilde{\chi}$ is a good coloring with $d_n(\chi, \widetilde{\chi}) = \delta$ then $1 - \widetilde{\chi}$ is a good coloring with $d_n(\chi, 1 - \widetilde{\chi}) = 1 - \delta$. Note that

$$\mathbb{E}_n^u\left(\sum_{0\leq\delta\leq 2^{-k/2}} Z_g(\delta)\Big|\chi \in S_g\right) \leq \mathbb{E}_n^u\big(\#\mathcal{C}_\sigma(\chi)|\chi \in S_g\big) \leq \mathbb{E}_n^u(Z_e) \quad (3.25)$$

where the last inequality holds by definition of S_g.

For colorings not in the local cluster,

$$\begin{aligned}\mathbb{E}_n^u\left(\sum_{2^{-k/2}<\delta\leq 1/2} Z_g(\delta)\Big|\chi \in S_g\right) &\leq \mathbb{E}_n^u\left(\sum_{2^{-k/2}<\delta\leq 1/2} Z_e(\delta)\Big|\chi \in S_g\right) \\ &\leq \mathbb{E}_n^u\left(\sum_{2^{-k/2}<\delta\leq 1/2} Z_e(\delta)\Big|\chi \text{ proper}\right)\frac{\mathbb{P}_n^u(\chi \text{ proper})}{\mathbb{P}_n^u(\chi \in S_g)}\end{aligned}$$

where the sum is over all $\delta \in \mathbb{Z}[1/n]$ in the given range. By definition and Proposition 3.4.11,

$$\frac{\mathbb{P}_n^u(\chi \text{ proper})}{\mathbb{P}_n^u(\chi \in S_g)} = \frac{1}{\mathbb{P}_n^u(\chi \in S_g|\chi \text{ proper})} = \frac{1}{\mathbb{P}_n^\chi\big(|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)\big)} \to 1$$

as $n \to \infty$. Since $\mathbb{E}_n^u(\cdot|\chi \text{ proper}) = \mathbb{E}_n^\chi(\cdot)$, the above inequality now implies

$$\begin{aligned}\mathbb{E}_n^u\left(\sum_{2^{-k/2}<\delta\leq 1/2} Z_g(\delta)\Big|\chi \in S_g\right) &\lesssim \sum_{2^{-k/2}<\delta\leq 1/2} \mathbb{E}_n^\chi(Z_e(\delta)) \leq C_1 \sum_{2^{-k/2}<\delta\leq 1/2} e^{n\psi_0(\delta)} \\ &\leq C_1 n e^{nf(d,k)} \leq C_2\mathbb{E}_n^u(Z_e) \quad (3.26)\end{aligned}$$

where the second inequality holds by Theorem 3.4.1 for some function $C_1 = C_1(n,k,r)$ which is sub-exponential in n. The second-to-last inequality holds because the number of summands is bounded by n since δ is constrained to lie in $\mathbb{Z}[1/n]$ and by Proposition 3.4.10, $\psi_0(\delta) \leq f(d,k)$. The last inequality holds for some function $C_2 = C_2(n,k,r)$ that is sub-exponential in n since by Theorem 3.3.2, $n^{-1}\log \mathbb{E}_n^u(Z_e)$ converges to $f(d,k)$.

Combine (3.24), (3.25) and (3.26) to obtain

$$\mathbb{E}_n^u(Z_g|\chi \in S_g) \leq 2(1+C_2)\mathbb{E}_n^u(Z_e).$$

Plug this into (3.23) to obtain

$$\mathbb{E}_n^u(Z_g^2) \leq 2(1+C_2)\mathbb{E}_n^u(Z_e)^2 \sim 2(1+C_2)\mathbb{E}_n^u(Z_g)^2$$

where the asymptotic $\sim$ holds by Lemma 3.4.12. This proves the lemma.

□

Corollary 3.4.14. *Lemma 3.0.6 inequality (3.3) is true. That is:*

$$0 \;=\; \inf_{\epsilon>0} \liminf_{n\to\infty} n^{-1}\log\left(\mathbb{P}_n^u\left(\left|n^{-1}\log Z(\sigma) - f(d,k)\right| < \epsilon\right)\right).$$

Proof. By Theorem 3.3.2,

$$n^{-1}\log \mathbb{E}_n^u(Z(\sigma)) \to f(d,k)$$

as $n\to\infty$. In particular, for every $\epsilon>0$, for large enough n, $\mathbb{P}_n^u\left(n^{-1}\log Z(\sigma) > f(d,k)+\epsilon\right) < \exp\left(-\frac{\epsilon}{2}n\right)$. So it suffices to prove

$$0 \;=\; \inf_{\epsilon>0} \liminf_{n\to\infty} n^{-1}\log\left(\mathbb{P}_n^u\left(n^{-1}\log Z(\sigma) \geq f(d,k)-\epsilon\right)\right).$$

Since $Z(\sigma) \geq Z_g(\sigma)$, it suffices to prove the same statement with $Z_g(\sigma)$ in place of $Z(\sigma)$. By Lemma 3.4.12 and Theorem 3.3.2, $n^{-1}\log\left(\mathbb{E}_n^u[Z_g(\sigma)]\right)$ converges to $f(d,k)$ as $n \to \infty$. So we may replace $f(d,k)$ in the statement above with $n^{-1}\log\left(\mathbb{E}_n^u[Z_g(\sigma)]\right)$. Then we may multiply by n both sides and exponentiate inside the probability. So it suffices to prove

$$0 \;=\; \inf_{\epsilon>0} \liminf_{n\to\infty} n^{-1}\log\left(\mathbb{P}_n^u\left(Z_g(\sigma) \geq \mathbb{E}_n^u[Z_g(\sigma)]e^{-n\epsilon}\right)\right). \tag{3.27}$$

By the Paley-Zygmund inequality and Lemma 3.4.13

$$\mathbb{P}_n^u\left(Z_g(\sigma) > \mathbb{E}_n^u[Z_g(\sigma)]e^{-n\epsilon}\right) \geq (1-e^{-n\epsilon})^2\frac{\mathbb{E}_n^u[Z_g(\sigma)]^2}{\mathbb{E}_n^u[Z_g(\sigma)^2]} \geq \frac{1}{C}$$

where $C = C(n)$ is sub-exponential in n. This implies (3.27). □

3.5 The local cluster

To prove Proposition 3.4.11, we show that with high probability in the planted model, there is a 'rigid' set of vertices with density approximately $1 - 2^{-k}$. Rigidity here means that any proper coloring either mostly agrees with the planted coloring on the rigid set or it must disagree on a large density subset. Before making these notions precise, we introduce the various subsets, state precise lemmas about them and prove Proposition 3.4.11 from these lemmas which are proven in the next two sections.

So suppose $G = (V, E)$ is a k-uniform d-regular hyper-graph and $\chi : V \to \{0, 1\}$ is a proper coloring. An edge $e \in E$ is **χ-critical** if there is a vertex $v \in e$ such that $\chi(v) \notin \chi(e \setminus \{v\})$. If this is the case, then we say v **supports** e **with respect to** χ. If χ is understood then we will omit mention of it. We will apply these notions both to the case when G is the Cayley hyper-tree of Γ and when $G = G_\sigma$ is a finite hyper-graph.

For $l \in \{0, 1, 2, \ldots\}$, define the **depth l-core of** χ to be the subset $C_l(\chi) \subset V$ satisfying

$$C_0(\chi) = V,$$

$$C_{l+1}(\chi) = \{v \in V : \text{ } v \text{ supports at least 3 edges } e \text{ such that } e \setminus \{v\} \subset C_l(\chi)\}.$$

By induction, $C_{l+1}(\chi) \subset C_l(\chi)$ for all $l \geq 0$. Also let $C_\infty(\chi) = \cap_l C_l(\chi)$.

The set $C_l(\chi)$ is defined to consist of vertices v so that if v is re-colored (in some proper coloring) then this re-coloring forces a sequence of re-colorings in the shape of an immersed hyper-tree of degree at least 3 and depth l. Re-coloring a vertex of $C_\infty(\chi)$ would force re-coloring an infinite immersed tree of degree at least 3.

Also define the **attached vertices** $A_l(\chi) \subset V$ by: $v \in A_l(\chi)$ if $v \notin C_l(\chi)$ but there exists an edge e, supported by v such that $e \setminus \{v\} \subset C_{l-1}(\chi)$. Thus if $v \in A_l(\chi)$ is re-colored then it forces a re-coloring of some vertex in $C_{l-1}(\chi)$. In this definition, we allow $l = \infty$ (letting $\infty - 1 = \infty$).

In order to avoid over-counting, we also need to define the subset $A'_l(\chi)$ of vertices $v \in A_l(\chi)$ such that there exists a vertex $w \in A_l(\chi)$, with $w \neq v$, and edges e_v, e_w supported by v, w respectively such that

1. $e_v \setminus \{v\}, e_w \setminus \{w\} \subset C_{l-1}(\chi)$,
2. $e_v \cap e_w \neq \emptyset$.

In this definition, we allow $l = \infty$.

We will need the following constants:

$$\lambda_0 = \frac{1}{2^{k-1} - 1}, \quad \lambda := d\lambda_0 = \frac{d}{2^{k-1} - 1}.$$

The significance of λ_0 is: if e is an edge and $v \in e$ a vertex then λ_0 is the probability v supports e in a uniformly random proper coloring of e. So $\lambda = d\lambda_0$ is the expected

number of edges that v supports. For the values of d and k used in the Key Lemma 3.0.6, $\lambda = \log(2)k + O(k2^{-k})$.

For the next two lemmas, we assume the hypotheses of Proposition 3.4.11.

Lemma 3.5.1. *For any $\delta > 0$ there exists k_0 such that $k \geq k_0$ implies*

$$\lim_{l\to\infty} \liminf_{n\to\infty} \mathbb{P}_n^{\chi} \left(\frac{|C_l(\chi) \cup A_l(\chi) \setminus A_l'(\chi)|}{n} > 1 - e^{-\lambda}(1+\delta) \right) = 1.$$

Lemma 3.5.1 is proven in §3.6.

Definition 3.5.2. *Fix a proper 2-coloring $\chi : V \to \{0,1\}$. Let $\rho > 0$. A subset $R \subset V$ is* **ρ-rigid** *(with respect to χ) if for every proper coloring $\chi' : V \to \{0,1\}$, $|\{v \in R : \chi(v) \neq \chi'(v)\}|$ is either less than $\rho|V|$ or greater than $2^{-k/2}|V|$.*

Lemma 3.5.3. *For any $\rho > 0$,*

$$\lim_{l\to\infty} \liminf_{n\to\infty} \mathbb{P}_n^{\chi} \left(C_l(\chi) \cup A_l(\chi) \setminus A_l'(\chi) \text{ is } \rho\text{-rigid} \right) = 1.$$

Lemma 3.5.3 is proven in §3.7. We can now prove Proposition 3.4.11:

Proposition 3.4.11. *Let $0 < \eta_0 < \eta_1 < (1 - \log 2)/2$. Then for all sufficiently large k (depending on η_0, η_1), if*

$$r := d/k = \frac{\log(2)}{2} \cdot 2^k - (1 + \log(2))/2 + \eta$$

for some η with $\eta_0 \leq \eta \leq \eta_1$ then with high probability in the planted model, $|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)$. In symbols,

$$\lim_{n\to\infty} \mathbb{P}_n^{\chi}\big(|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)\big) = 1.$$

Proof. Let $0 < \rho, \delta$ be small constants satisfying

$$\log(2)\delta + H(\rho, 1-\rho) + \log(2)\rho < (1 - 2\eta - \log(2))2^{-k}. \tag{3.28}$$

Let l be a natural number. Also let $\sigma : \Gamma \to \mathrm{Sym}(n)$ be a uniform homomorphism and $\chi : [n] \to \{0,1\}$ a proper coloring. To simplify notation, let

$$R = C_l(\chi) \cup A_l(\chi) \setminus A'_l(\chi).$$

By Lemmas 3.5.1 and 3.5.3 it suffices to show that if $|R|/n > 1 - e^{-\lambda} - \delta$ and R is ρ-rigid then $|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}^u_n(Z_e)$ (for all sufficiently large n). So assume $|R|/n > 1-e^{-\lambda}-\delta$ and R is ρ-rigid.

Let $\chi' \in \mathcal{C}_\sigma(\chi)$. By definition, this means $d_n(\chi', \chi) \leq 2^{-k/2}$. Since R is ρ-rigid, this implies

$$|\{v \in R : \chi(v) \neq \chi'(v)\}| \leq \rho n.$$

Since this holds for all $\chi' \in \mathcal{C}_\sigma(\chi)$, it follows that

$$|\mathcal{C}_\sigma(\chi)| \leq \binom{|R|}{\rho n} 2^{\rho n} 2^{n-|R|}.$$

By Stirling's formula

$$n^{-1} \log \binom{|R|}{\rho n} \leq n^{-1} \log \binom{n}{\rho n} \leq H(\rho, 1-\rho) + O(n^{-1}\log(n)).$$

Since $|R|/n > 1 - e^{-\lambda} - \delta = 1 - 2^{-k} - \delta + O(k2^{-2k})$,

$$n^{-1} \log(2^{n-|R|}) \leq \log(2)[2^{-k} + \delta] + O(k2^{-2k}).$$

Thus,

$$n^{-1} \log |\mathcal{C}_\sigma(\chi)| \leq \log(2)2^{-k} + \log(2)\delta + H(\rho, 1-\rho) + \log(2)\rho + O(k2^{-2k} + n^{-1}\log(n)).$$

On the other hand,

$$n^{-1}\log \mathbb{E}_n^u(Z_e) = f(d,k) + O(n^{-1}\log(n)) = (1-2\eta)2^{-k} + O(2^{-2k}) + O(n^{-1}\log(n))$$

by Lemma 3.4.8 and Theorem 3.3.2.

Therefore, the choice of ρ, δ in (3.28) implies $|\mathcal{C}_\sigma(\chi)| \leq \mathbb{E}_n^u(Z_e)$ for all sufficiently large n. This also depends on k being sufficiently large, but the lower bound on k is uniform in n.

□

3.6 A Markov process on the Cayley hyper-tree

Here we study a Γ-invariant measure μ on X which is, in a sense, the limit of the planted model. We will define it by specifying its values on cylinder sets.

Let $D \subset \Gamma$ be a connected finite union of hyper-edges. Let $\xi \in \{0,1\}^D$ be a proper coloring. We define $C(\xi) \subset X$ to be the set of proper colorings $\xi' \in X$ with $\xi' \restriction D = \xi$ (where $\restriction$ means "restricted to"). We set $\mu(C(\xi))$ equal to the reciprocal of the number of proper colorings of D. In particular, if $\xi' \in \{0,1\}^D$ is another proper coloring of D, then $\mu(C(\xi)) = \mu(C(\xi'))$.

Because X is totally disconnected, any Borel probability measure on X is determined by its values on clopen subsets (since these generate the topology). Since every clopen subset is a finite union of cylinder sets of the form above, Kolmogorov's Extension Theorem implies the existence of a unique Borel probability μ on X satisfying the aforementioned equalities.

Note this measure has the following Markov property. Let $x = (x_g)_{g\in G}$ be a random element of X with law μ. Let $v \in \Gamma$ and let $e \subset \Gamma$ be a hyperedge containing v. Let $\mathrm{Past}(e,v)$ be the set of all $g \in \Gamma$ such that every path in the Cayley hyper-tree from g to an element of e passes through v. In particular, $e \cap \mathrm{Past}(e,v) = \{v\}$.

Then the distribution of $(x_g)_{g\in e\setminus\{v\}}$ conditioned on $\{x_g :\ g \in \mathrm{Past}(e,v)\}$ is uniformly distributed on the set of all colorings $y : (e \setminus \{v\}) \to \{0,1\}$ such that there exists some $h \in e \setminus \{v\}$ with $y(h) \neq x(v)$.

3.6.1 Local convergence

We will prove the following lemma.

Lemma 3.6.1. *Let $\chi : V \to \{0,1\}$ be an equitable coloring with $|V| = n$. If $B \subseteq X$ is clopen, then for every $\epsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}_n^{\chi}\left(\left|\frac{1}{n}\sum_{v\in V} \mathbb{1}_B(\Pi_v^{\sigma_n}(\chi)) - \mu(B)\right| > \epsilon\right) = 0.$$

To prove this lemma we will first show that if $f_n : \mathrm{Hom}_{\chi}(\Gamma, \mathrm{Sym}(n)) \to \mathbb{R}$ is the function

$$f_n(\sigma_n) := \frac{1}{n}\sum_{v\in V} \mathbb{1}_B(\Pi_v^{\sigma_n}(\chi))$$

then f_n concentrates about its expectation using Theorem B.0.2, and then we will show that this expectation is given by $\mu(B)$.

Proposition 3.6.2. *We have*

$$\lim_{n\to\infty} \mathbb{P}_n^{\chi}\left(|f_n - \mathbb{E}_n^{\chi}[f_n]| > \epsilon\right) = 0.$$

Proof. For $g \in \Gamma$, let $\mathrm{pr}_g : X \to \{0,1\}$ be the projection map $\mathrm{pr}_g(x) = x_g$. For $D \subset \Gamma$, let $\mathcal{F}_D$ be the smallest Borel sigma-algebra such that pr_g is $\mathcal{F}_D$-measurable for every $g \in D$.

Since every clopen subset B of X is a finite union of cylinder sets, the function $\mathbb{1}_B$ is $\mathcal{F}_D$-measurable for some finite set $D \subset \Gamma$.

We will use the normalized Hamming metrics $d_{\mathrm{Sym}(n)}$ and d_{Hom} on $\mathrm{Sym}(n)$ and $\mathrm{Hom}_{\chi}(\Gamma, \mathrm{Sym}(n))$ respectively. These are defined in the beginning of Appendix B.

We claim f_n is L-Lipschitz for some $L < \infty$. Let $\sigma, \sigma' \in \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$. Because $\mathbb{1}_B$ is $\mathcal{F}_D$-measurable,

$$\begin{aligned}|f_n(\sigma) - f_n(\sigma')| &\leq n^{-1}\#\{v \in [n] : \chi(\sigma(\gamma^{-1})(v)) \neq \chi(\sigma'(\gamma^{-1})(v)) \text{ for some } \gamma \in D\} \\ &\leq n^{-1}\#\{v \in [n] : \sigma(\gamma^{-1})(v) \neq \sigma'(\gamma^{-1})(v) \text{ for some } \gamma \in D\} \\ &\leq \sum_{\gamma \in D} d_{\mathrm{Sym}(n)}(\sigma(\gamma^{-1}), \sigma'(\gamma^{-1})).\end{aligned}$$

Now $d_{\mathrm{Sym}(n)}$ is both left and right invariant. So

$$d_{\mathrm{Sym}(n)}(gh, g'h') \leq d_{\mathrm{Sym}(n)}(gh, gh') + d_{\mathrm{Sym}(n)}(gh', g'h') = d_{\mathrm{Sym}(n)}(h, h') + d_{\mathrm{Sym}(n)}(g, g')$$

for any $g, g', h, h' \in \mathrm{Sym}(n)$. Furthermore we immediately have for any $1 \leq i \leq d$ that $d_{\mathrm{Sym}(n)}(\sigma(s_i), \sigma'(s_i)) \leq d_{\mathrm{Hom}}(\sigma, \sigma')$. Together these imply $d_{\mathrm{Sym}(n)}(\sigma(\gamma), \sigma'(\gamma)) \leq |\gamma| d_{\mathrm{Hom}}(\sigma, \sigma')$ for any $\gamma \in \Gamma$ where $|\gamma|$ is the distance from γ to the identity in the word metric on Γ. Thus if we take $L = \sum_{\gamma \in D} |\gamma| < \infty$ we see that $|f_n(\sigma) - f_n(\sigma')| \leq L d_{\mathrm{Hom}}(\sigma, \sigma')$ as desired.

The Proposition now follows from Theorem B.0.2. □

To finish the proof of Lemma 3.6.1, it now suffices to show the expectation of f_n with respect to the planted model converges to $\mu(B)$ as $n \to \infty$. We will prove this by an inductive argument, the inductive step of which is covered in the next lemma.

Lemma 3.6.3. *Let $D \subset \Gamma$ be either a singleton or a connected finite union of hyperedges containing the identity. Let $\xi \in \{0,1\}^D$ be a proper coloring. Let $F_{D,\xi,v} \subset \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ be the event that $\Pi_v^{\sigma_n}(\chi) \restriction D = \xi \restriction D$. Let $Q(D)$ be the number of proper colorings of D. Then*

$$\lim_{n \to \infty} \sup_{v \in \chi^{-1}(\xi(1_\Gamma))} \left| \mathbb{P}_n^\chi \left(F_{D,\xi,v} \right) - \frac{2}{Q(D)} \right| = 0. \tag{3.29}$$

Remark 3.6.4. *The factor of* 2 *in* $\frac{2}{Q(D)}$ *is there to account for the fact that we are requiring* $v \in \chi^{-1}(\xi(1_\Gamma))$.

Proof. The statement is immediate if $D = \{1_\Gamma\}$ is a singleton. For induction we assume that the statement is true for some finite $D \subset \Gamma$. Let $D' \supset D$ be a connected union of hyper-edges such that there exists a unique hyper-edge e with $D' = D \cup e$. By induction, it suffices to prove the statement for D'.

Note by symmetry that $\mathbb{P}_n^\chi(F_{D,\xi,v})$ is the same for all $v \in \chi^{-1}(\xi(1_\Gamma))$. Let us denote this common value by $c_{D,\xi}$. Define $A_{D,\xi} = \{(v,\sigma) : v \in V, \sigma \in F_{D,\xi,v}\}$. If $g \in D$ then $(v,\sigma) \mapsto (\sigma(g^{-1})v, \sigma)$ is a bijection from $A_{D,\xi}$ to $A_{g^{-1}D,g^{-1}\xi}$. This implies

$$\begin{aligned}\frac{n}{2} c_{D,\xi} \left|\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))\right| &= |A_{D,\xi}| \\ &= |A_{g^{-1}D,g^{-1}\xi}| \\ &= \frac{n}{2} c_{g^{-1}D,g^{-1}\xi} \left|\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))\right|.\end{aligned}$$

Thus $c_{g^{-1}D,g^{-1}\xi} = c_{D,\xi}$. After replacing D' with $\gamma^{-1}D'$ where $\gamma \in e \cap D$ is the unique element in $e \cap D$, we may therefore assume without loss of generality that $e \cap D = \{1_\Gamma\}$. By symmetry we may also assume $\xi(1_\Gamma) = 1$ and $e = \{s_1^i : \ i = 0, \cdots, k-1\}$ is the hyper-edge generated by s_1.

For any subset $W \subset \Gamma$, let $E_{W,v}$ be the event that $\sigma_n(g^{-1})(v) \neq \sigma_n(g'^{-1})(v)$ for any $g, g' \in W$ with $g \neq g'$. Since the planted model is a random sofic approximation, by Lemma 3.2.1 we have $\mathbb{P}_n^\chi(E_{D\cup e,v}) = 1 - o_n(1)$ uniformly in v. We will show that

$$\lim_{n\to\infty} \sup_{v \in \chi^{-1}(1)} \left| \mathbb{P}_n^\chi \left(F_{e,\xi,v} \middle| F_{D,\xi,v}, E_{D\cup e,v} \right) - \frac{1}{2^{k-1}-1} \right| = 0. \tag{3.30}$$

Next we show how to finish the lemma assuming (3.30). We claim that $Q(D \cup e) = Q(D)(2^{k-1}-1)$. To see this, observe that if ξ' is a proper coloring of D and ξ'' is any coloring of $e \setminus \{1_\Gamma\}$ then the concatenation $\xi' \sqcup \xi''$ is a proper coloring of $D \cup e$

unless $\xi''(g) = \xi'(1_\Gamma)$ for all $g \in e \setminus \{1_\Gamma\}$. Since $|e \setminus \{1_\Gamma\}| = k-1$, this implies that every proper coloring of D admits $2^{k-1}-1$ proper extensions to $D \cup e$. So $Q(D \cup e) = Q(D)(2^{k-1}-1)$ as claimed.

Since $F_{D\cup e,\xi,v} = F_{e,\xi,v} \cap F_{D,\xi,v}$,

$$\begin{aligned}\mathbb{P}_n^\chi\left(F_{D\cup e,\xi,v}\right) &= \mathbb{P}_n^\chi\left(F_{e,\xi,v}\middle| F_{D,\xi,v}\right)\mathbb{P}_n^\chi\left(F_{D,\xi,v}\right)\\ &= \mathbb{P}_n^\chi\left(F_{e,\xi,v}\middle| F_{D,\xi,v}, E_{D\cup e,v}\right)\mathbb{P}_n^\chi\left(F_{D,\xi,v}\right)+o_n(1).\end{aligned}$$

The lemma now follows from the inductive hypothesis, (3.30) and $Q(D \cup e) = Q(D)(2^{k-1}-1)$.

It now suffices to prove (3.30). Let D_j denote the union of all subsets $\{gs_j^i : i = 0, \cdots, k-1\}$ over all $g \in D$ such that $\{gs_j^i : i = 0, \cdots, k-1\} \subseteq D$. That is, D_j is the union of all hyperedges generated by s_j which lie completely within D. Because $e = \{s_1^i : \ i = 0, \cdots, k-1\}$, $e \cap D_1 = \emptyset$. Multiplication by s_j on the right preserves D_j.

Fix $v \in V = \{1, \ldots, n\}$ and consider an injective function $h : D \to V$ such that $h(1_\Gamma) = v$ and $\chi\left(h\left(e'\right)\right) = \{0,1\}$ for all hyper-edges $e' \in D$. Let $N(\chi, h)$ be the set of $\sigma \in \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ such that $\sigma(g^{-1})v = h(g)$ for all $g \in D$. By definition,

$$\mathbb{P}_n^\chi\left(\sigma_n(g^{-1})(v) = h(g)\ \forall g \in D\right) = \frac{|N(\chi,h)|}{|\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))|}.$$

For each $j \in \{1, \ldots, d\}$, let $N_j(\chi, h)$ be the set of permutations $\pi_j \in \mathrm{Sym}(n)$ such that

1. the orbit-partition of π_j is a k-partition which is properly colored by χ,

2. $h(gs_j) = \pi_j^{-1}h(g)$ for all $g \in D \cap Ds_j^{-1}$.

We claim that the map

$$N(\chi, h) \mapsto N_1(\chi, h) \times \cdots \times N_d(\chi, h)$$

that sends σ to $(\sigma(s_1), \ldots, \sigma(s_d))$ is a bijection.

The fact that it is injective is immediate since Γ is generated by $s_1, \ldots, s_d$. In order to prove that it is surjective, fix elements $\pi_j \in N_j(\chi, h)$ for all j. Define $\sigma \in \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ by $\sigma(s_j) = \pi_j$. It suffices to check that $\sigma(g^{-1})(v) = h(g)\ \forall g \in D$. By induction on the number of edges in D, it suffices to assume that the equation $\sigma(g^{-1})(v) = h(g)$ holds for some $g \in D \cap Ds_j^{-1}$ and prove $\sigma((gs_j)^{-1})(v) = h(gs_j)$. This follows from:

$$\sigma((gs_j)^{-1})(v) = \sigma(s_j^{-1})\sigma(g^{-1})(v) = \pi_j^{-1}h(g) = h(gs_j).$$

This proves the claim. Therefore,

$$\mathbb{P}_n^\chi\left(\sigma_n(\cdot)^{-1}(v) \restriction D = h\right) = \frac{\prod_{j=1}^d |N_j(\chi, h)|}{|\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))|}. \tag{3.31}$$

Let $H_D = \{h : D \to V : h \text{ is injective, } \chi(h(g)) = \xi(g)\ \forall g \in D, \text{ and } h(1_\Gamma) = v\}$. We have

$$\mathbb{P}_n^\chi(F_{D,\xi,v}, E_{D,v}) = \sum_{h \in H_D} \mathbb{P}_n^\chi(\sigma_n(\cdot)^{-1}(v) \restriction D = h).$$

Similarly

$$\mathbb{P}_n^\chi(F_{D\cup e,\xi,v}, E_{D\cup e,v}) = \sum_{h' \in H_{D\cup e}} \mathbb{P}_n^\chi(\sigma_n(\cdot)^{-1}(v) \restriction D \cup e = h').$$

Each $h \in H_D$ is the restriction to D of

$$(m-1)!\binom{n/2 - \#\xi^{-1}(1) \cap D}{m-1}(k-m)!\binom{n/2 - \#\xi^{-1}(0) \cap D}{k-m} \tag{3.32}$$

distinct $h' \in H_{D \cup e}$ where $m = |\xi^{-1}(1) \cap e|$. We can express

$$\begin{aligned}
\mathbb{P}_n^\chi(F_{e,\xi,v} | F_{D,\xi,v}, E_{D\cup e,v}) &= \frac{\mathbb{P}_n^\chi(F_{D\cup e,\xi,v}, E_{D\cup e,v})}{\mathbb{P}_n^\chi(F_{D,\xi,v}, E_{D\cup e,v})} \\
&= \frac{\mathbb{P}_n^\chi(F_{D\cup e,\xi,v}, E_{D\cup e,v})}{\mathbb{P}_n^\chi(F_{D,\xi,v}, E_{D,v})} + o_n(1) \\
&= \frac{\sum_{h' \in H_{D\cup e}} \mathbb{P}_n^\chi(\sigma_n(\cdot)^{-1}(v) \restriction D \cup e = h')}{\sum_{h \in H_D} \mathbb{P}_n^\chi(\sigma_n(\cdot)^{-1}(v) \restriction D = h)} + o_n(1) \\
&= \frac{\sum_{h' \in H_{D\cup e}} \prod_{j=1}^d |N_j(\chi, h')|}{\sum_{h \in H_D} \prod_{j=1}^d |N_j(\chi, h)|} + o_n(1)
\end{aligned}$$

where the $o_n(1)$ term follows from the inductive hypothesis and the fact that $\mathbb{P}_n^\chi(E_{D,v}) \geq \mathbb{P}_n^\chi(E_{D\cup e,v}) = 1 - o_n(1)$.

If $h_1, h_2 \in H_D$ then $\chi(h_1(g)) = \xi(g) = \chi(h_2(g))$ for all $g \in D$. Observe that $|N_j(\chi, h_1)| = |N_j(\chi, h_2)|$. This follows by conjugating by a permutation in $\mathrm{Sym}(n,k)$ which preserves the color classes of χ and which maps $h_1(g)$ to $h_2(g)$ for each $g \in D$. Similarly, if $h_1', h_2' \in H_{D\cup e}$ then $|N_j(\chi, h_1')| = |N_j(\chi, h_2')|$. Therefore,

$$\mathbb{P}_n^\chi(F_{e,\xi,v} | F_{D,\xi,v}, E_{D\cup e,v}) = \frac{|H_{D\cup e}| \prod_{j=1}^d |N_j(\chi, h')|}{|H_D| \prod_{j=1}^d |N_j(\chi, h)|} + o_n(1)$$

for any $h \in H_D$ and $h' \in H_{D\cup e}$.

Next we choose $h' \in H_{D\cup e}$ and $h \in H_D$ so that h' extends h. Note $|N_j(\chi, h')| =$

$|N_j(\chi,h)|$ for all $j \neq 1$ because $D_j = (D \cup e)_j$. Therefore,

$$\begin{aligned}
&\mathbb{P}_n^\chi(F_{e,\xi,v}|F_{D,\xi,v}, E_{D\cup e,v}) \\
&= \frac{|H_{D\cup e}||N_1(\chi,h')|}{|H_D||N_1(\chi,h)|} + o_n(1) \\
&= \frac{|N_1(\chi,h')|}{|N_1(\chi,h)|}(m-1)!\binom{n/2 - \#\xi^{-1}(1)\cap D}{m-1}(k-m)!\binom{n/2-\#\xi^{-1}(0)\cap D}{k-m} + o_n(1) \\
&\sim \frac{|N_1(\chi,h')|}{|N_1(\chi,h)|}(n/2)^{k-1} \qquad (3.33)
\end{aligned}$$

where the second equality uses (3.32). Here we are using the notation $f(n) \sim g(n)$ to mean $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$.

To compute $|N_1(\chi,h)|$, let

$$T_{D,n} = \left\{ \vec{t} \in [0,1]^{k+1} : t_0 = t_k = 0, \sum_i t_i = \frac{1}{k}, \sum_i i t_i = \frac{\#\chi^{-1}(1)\setminus h(D_1)}{\#V \setminus h(D_1)}, (n-\#D_1)t_i \in \mathbb{Z} \right\}$$

be the set of possible types of orbit-partitions of permutations of $V \setminus h(D_1)$ that contribute to the count $|N_1(\chi,h)|$. To be precise, if $\vec{t} \in T_{D,n}$ then there is a k-partition of $V \setminus h(D_1)$ such that the number of parts P of the partition with $|P \cap \chi^{-1}(1)| = i$ is $t_i(n - \#D_1)$.

If we fix $\vec{t} \in T_{D,n}$ then the number of permutations whose orbit partition has type $\vec{t}$ is

$$(k-1)!^{(n-\#D_1)/k} \frac{(n/2-\#\xi^{-1}(1)\cap D_1)!(n/2-\#\xi^{-1}(0)\cap D_1)!}{\prod_{j=0}^k j!^{t_j(n-\#D_1)}(k-j)!^{t_j(n-\#D_1)}(t_j(n-\#D_1))!}$$

where we have used (3.8). It follows that

$$|N_1(\chi,h)| = (k-1)!^{(n-\#D_1)/k} \frac{(n/2-\#\xi^{-1}(1)\cap D_1)!(n/2-\#\xi^{-1}(0)\cap D_1)!}{\left(\frac{n-\#D_1}{k}\right)!} \times \sum_{\vec{t}\in T_{D,n}} S_{D,n}(\vec{t})$$

where

$$S_{D,n}(\vec{t}) = \frac{((n-\#D_1)/k)!}{\prod_{i=1}^{k-1} i!^{t_i(n-\#D_1)}(k-i)!^{t_i(n-\#D_1)}(t_i(n-\#D_1))!}. \tag{3.34}$$

So

$$\frac{|N_1(\chi,h')|}{|N_1(\chi,h)|} \sim \frac{(n/k)}{(k-1)!(n/2)^k} \frac{\sum_{\vec{t}\in T_{D\cup e,n}} S_{D\cup e,n}(\vec{t})}{\sum_{\vec{t}\in T_{D,n}} S_{D,n}(\vec{t})}.$$

We plug this into (3.33) to obtain

$$\lim_{n\to\infty} \mathbb{P}_n^{\chi}(F_{e,\xi,v}|F_{D,\xi,v}, E_{D\cup e,v}) = \lim_{n\to\infty} \frac{2}{k!} \frac{\sum_{\vec{t}\in T_{D\cup e,n}} S_{D\cup e,n}(\vec{t})}{\sum_{\vec{t}\in T_{D,n}} S_{D,n}(\vec{t})}. \tag{3.35}$$

Define $t^* \in [0,1]^{k+1}$ by $t_0^* = t_k^* = 0$ and $t_i^* = \frac{1}{k(2^k-2)}\binom{k}{i}$ for $0 < i < k$. We establish asymptotic estimates of the sum of $S_{D,n}(\vec{t})$ in Lemma 3.6.5 and show that it suffices to sum over a small ball around t^*.

Recall $m = |\xi^{-1}(1) \cap e|$. Let $\tilde{T}_{D,n} = \{\vec{t} \in T_{D,n} : t_m > 0\}$. We define $f_n : \tilde{T}_{D,n} \to T_{D\cup e,n}$ by

$$f_n(\vec{t})_j = \begin{cases} \frac{(n-\#D_1)t_j - 1}{n-\#D_1-k} \text{ if } j = m \\ \frac{n-\#D_1}{n-\#D_1-k} t_j \text{ if } j \neq m \end{cases}.$$

Because $e \cap D_1 = \emptyset$, if $\vec{t} \in \tilde{T}_{D,n}$ then $f_n(\vec{t}) \in T_{D\cup e,n}$.

We claim that for any $\delta > 0$ there exists an $N := N(\delta)$ and $\delta' > 0$ such that for $n > N$, $T_{D\cup e,n} \cap B_{\delta'}(t^*) \subseteq f_n(\tilde{T}_{D,n} \cap B_\delta(t^*))$. This follows from observing that f_n is invertible and that $d(f_n(\vec{t}), \vec{t}) \leq \frac{k^2}{n-\#D_1-k}$. By Lemma 3.6.5 we have

$$\lim_{n\to\infty} \frac{\sum_{\vec{t}\in T_{D\cup e,n}} S_{D\cup e,n}(\vec{t})}{\sum_{\vec{t}\in T_{D,n}} S_{D,n}(\vec{t})} = \lim_{n\to\infty} \frac{\sum_{\vec{t}\in \tilde{T}_{D,n}\cap B_\delta(t^*)} S_{D\cup e,n}(f_n(\vec{t}))}{\sum_{\vec{t}\in T_{D,n}\cap B_\delta(t^*)} S_{D,n}(\vec{t})}. \tag{3.36}$$

Furthermore, since $1 \leq m \leq k-1$ and $t_m^* > 0$, for $\delta > 0$ sufficiently small we have $\tilde{T}_{D,n} \cap B_\delta(t^*) = T_{D,n} \cap B_\delta(t^*)$. Because $\#(D \cup e)_1 = \#D_1 + k$, the ratio $\frac{S_{D\cup e,n}(f_n(\vec{t}))}{S_{D,n}(\vec{t})}$ simplifies to

$$\frac{k}{n - \#D_1} m!(k-m)! t_m (n - \#D_1) = \frac{k(k!)}{\binom{k}{m}} t_m.$$

In particular, at t^* we obtain

$$\frac{S_{D\cup e,n}(f_n(t^*))}{S_{D,n}(t^*)} = \frac{k!}{2^k - 2}.$$

Suppose $\vec{t} \in B_\delta(t^*)$. Then $|t_m - t_m^*| < \delta$. So

$$\left| \frac{S_{D\cup e,n}(f_n(\vec{t}))}{S_{D,n}(\vec{t})} - \frac{k!}{2^k - 2} \right| = \left| \frac{S_{D\cup e,n}(f_n(\vec{t}))}{S_{D,n}(\vec{t})} - \frac{S_{D\cup e,n}(f_n(t^*))}{S_{D,n}(t^*)} \right| = \frac{k(k!)}{\binom{k}{m}} |t_m - t_m^*| \leq k! \delta.$$

Since δ is arbitrary, this and (3.36) imply

$$\lim_{n\to\infty} \frac{\sum_{\vec{t} \in T_{D\cup e,n}} S_{D\cup e,n}(\vec{t})}{\sum_{\vec{t} \in T_{D,n}} S_{D,n}(\vec{t})} = \frac{k!}{2^k - 2}.$$

We plug this into (3.35) to obtain

$$\lim_{n\to\infty} \mathbb{P}_n^\chi(F_{e,\xi,v} | F_{D,\xi,v}, E_{D\cup e,v}) = \frac{1}{2^{k-1} - 1}.$$

□

Lemma 3.6.5. *For any $\delta > 0$ we have*

$$\lim_{n\to\infty} \frac{\sum_{\vec{t} \in T_{D,n} \cap B_\delta(t^*)} S_{D,n}(\vec{t})}{\sum_{\vec{t} \in T_{D,n}} S_{D,n}(\vec{t})} = 1.$$

Proof. We use the following general consequence of Stirling's formula: let $l \in \mathbb{N}$, $\vec{t} \in [0,1]^l$ with $\sum_i t_i = 1$. Let $\vec{t}^{(n)}$ be any sequence in $[0,1]^l$ such that $n\vec{t}^{(n)} \in \mathbb{N}^l$ and

$\vec{t}^{(n)} \to \vec{t}$ as $n \to \infty$. If $\vec{t}^{(n)} = (t_{n,i})_{i=1}^{l}$ then $\lim_{n\to\infty} n^{-1} \log \binom{n}{nt_{n,1},\ldots,nt_{n,l}} = H(\vec{t})$.

By (3.34)

$$S_{D,n}(\vec{t}) = \binom{(n-\#D_1)/k}{t_{n,1}(n-\#D_1),\ldots,t_{n,k-1}(n-\#D_1)} \left(\prod_{i=1}^{k-1} \binom{k}{i}^{t_i(n-\#D_1)}\right) k!^{-(n-\#D_1)/k}.$$

By setting $N = (n-\#D_1)/k$, we obtain

$$\begin{aligned} & \lim_{n\to\infty} n^{-1} \log \binom{(n-\#D_1)/k}{t_{n,1}(n-\#D_1),\ldots,t_{n,k-1}(n-\#D_1)} \\ = \; & k^{-1} \lim_{n\to\infty} N^{-1} \log \binom{N}{t_{n,1}kN,\ldots,t_{n,k-1}kN} \\ = \; & k^{-1} H(k\vec{t}) = -\sum_{i=0}^{k} t_i \log(kt_i) = H(\vec{t}) - k^{-1}\log(k). \end{aligned}$$

We thus have:

$$\limsup_{n\to\infty} \sup_{\vec{t}\in T_{D,n}} \left| \frac{1}{n} \log S_{D,n}(\vec{t}) - \psi(\vec{t}) \right| = 0$$

where ψ is given by

$$\psi(\vec{t}) := -\frac{1}{k}\log(k(k)!) + H(\vec{t}) + \sum_{i=0}^{k} t_i \log \binom{k}{i}.$$

Note that ψ does not depend on n or D. Furthermore ψ is continuous and strictly concave. By the method of Lagrange multipliers (see Section 3.3.2) its maximum on $M := \{x \in [0,1]^{k+1} : x_0 = x_k = 0, \sum_i x_i = \frac{1}{k}\}$ occurs at t^*. In particular given $\delta > 0$ there is some $0 < \eta < \delta$ such that

$$\inf_{x\in M\cap B_\eta(t^*)} \psi(x) - \sup_{x\in M\setminus B_\delta(t^*)} \psi(x) > 0.$$

We claim that $B_\eta(t^*) \cap T_{D,n} \subseteq B_\eta(t^*) \cap M$ is non-empty for sufficiently large n. Since $\#T_{D,n} \leq n^k$ this claim and standard arguments involving the exponential growth of

$S_{D,n}(t)$ imply the lemma.

To prove the claim, we exhibit a member of $B_\eta(t^*) \cap T_{D,n}$ in a series of three steps. First, let $t^{(1)} \in \mathbb{R}^{k+1}$ be defined by $t_i^{(1)} = \frac{\lfloor t_i^*(n-\#D_1)\rfloor}{n-\#D_1}$ for each $0 \leq i \leq k$. This satisfies the integrality condition, that is $(n-\#D_1)t_i^{(1)} \in \mathbb{Z}$ for all i. Furthermore $|t^* - t^{(1)}| \leq \sum_{i=0}^k |t_i^* - t_i^{(1)}| \leq \frac{k-1}{n-\#D_1}$.

Second, define $t^{(2)} \in \mathbb{R}^{k+1}$ by

$$t_i^{(2)} = \begin{cases} t_1^{(1)} + \left(\frac{1}{k} - \sum_{j=1}^{k-1} t_j^{(1)}\right) & \text{if } i = 1 \\ t_i^{(1)} \text{ otherwise} \end{cases}.$$

Since $\frac{n-\#D_1}{k} \in \mathbb{Z}$ and $(n-\#D_1)t_i^{(1)} \in \mathbb{Z}$ for each i we maintain the integrality condition, and $\sum_i t_i^{(2)} = \frac{1}{k}$. Furthermore

$$\left|t^{(2)} - t^{(1)}\right| \leq \left|\frac{1}{k} - \sum_{i=1}^{k-1} t_i^{(1)}\right| = \left|\sum_{i=0}^{k}(t_i^* - t_i^{(1)})\right| \leq \sum_{i=0}^{k}\left|t_i^* - t_i^{(1)}\right| \leq \frac{k-1}{n-\#D_1}.$$

Finally let $\Delta = \frac{\#\chi^{-1}(1)\setminus h(D_1)}{n-\#D_1} - \sum_{i=0}^k i t_i^{(2)}$. Define $t^{(3)} \in \mathbb{R}^{k+1}$ by

$$t_i^{(3)} = \begin{cases} t_1^{(2)} - \Delta \text{ if } i = 1 \\ t_2^{(2)} + \Delta \text{ if } i = 2 \\ t_i^{(2)} \text{ otherwise} \end{cases}.$$

We claim $t^{(3)} \in B_\eta(t^*) \cap T_{D,n}$ for n sufficiently large. First $t_0^{(3)} = t_0^{(2)} = t_0^{(1)} = t_0^* = 0$, with the same equalities holding for $t_k^{(3)}$. We have $t_i^{(3)}(n-\#D_1) \in \mathbb{Z}$ since $t_i^{(2)}(n-\#D_1) \in \mathbb{Z}$ for each i and $\Delta(n-\#D_1) \in \mathbb{Z}$. We also have $\sum_i t_i^{(3)} = \sum_i t_i^{(2)} = \frac{1}{k}$. Furthermore $\sum_i i t_i^{(3)} = \sum_i i t_i^{(2)} + 2\Delta - \Delta = \frac{\chi^{-1}(1)\setminus h(D_1)}{n-\#D_1}$. Finally by repeated

application of the triangle inequality

$$|t^* - t^{(3)}| \leq \frac{k-1}{n - \#D_1} + \frac{k-1}{n - \#D_1} + 2|\Delta|.$$

We show that $|\Delta|$ is small for n sufficiently large. This will not only imply that $t^{(3)} \in B_\eta(t^*)$, but also that $t_i^{(3)} \in [0,1]$ for each i. Note that

$$\frac{n/2 - \#D_1}{n} \leq \frac{\#\chi^{-1}(1) \setminus h(D_1)}{n - \#D_1} \leq \frac{n/2}{n - \#D_1} = \frac{(n - \#D_1)/2 + \#D_1/2}{n - \#D_1}.$$

This implies $\left| \frac{\#\chi^{-1}(1) \setminus h(D_1)}{n - \#D_1} - \frac{1}{2} \right| \leq \max\{\frac{\#D_1}{n}, \frac{\#D_1}{2(n - \#D_1)}\}$. Thus

$$\begin{aligned}
|\Delta| &\leq \left| \frac{\#\chi^{-1}(1) \setminus h(D_1)}{n - \#D_1} - \frac{1}{2} \right| + \left| \frac{1}{2} - \sum_{i=0}^{k} i t_i^{(2)} \right| \\
&\leq \max\left\{ \frac{\#D_1}{n}, \frac{\#D_1}{2(n - \#D_1)} \right\} + \left| \sum_{i=0}^{k} i t_i^* - \sum_{i=0}^{k} i t_i^{(1)} \right| + \left| \sum_{i=0}^{k} i t_i^{(1)} - \sum_{i=0}^{k} i t_i^{(2)} \right| \\
&\leq \max\left\{ \frac{\#D_1}{n}, \frac{\#D_1}{2(n - \#D_1)} \right\} + \frac{2(k-1)^2}{n - \#D_1}.
\end{aligned}$$

As k and $\#D_1$ are fixed, we have $|t^* - t^{(3)}| < \eta$ for large enough n. Since $t_i^* \in (0,1)$ for each $1 \leq i \leq k-1$ we have $|t_i^{(3)} - t_i^*| \leq |t^{(3)} - t^*|$. Thus for sufficiently large n we will also have $t_i^{(3)} \in (0,1)$ for each $1 \leq i \leq k-1$. Therefore $t^{(3)} \in [0,1]^{k+1}$ and we have $t^{(3)} \in B_\eta(t^*) \cap T_{D,n}$ for n sufficiently large.

□

Proof of Lemma 3.6.1. Let μ_n^χ be the Borel probability measure on X defined by

$$\mu_n^\chi(B) = \mathbb{E}_n^\chi \left(\frac{1}{\#V} \sum_{v \in V} \mathbb{1}_B \left(\Pi_v^{\sigma_n}(\chi) \right) \right)$$

for any Borel set $B \subset X$. By Proposition 3.6.2, it suffices to show that $\mu_n^\chi(B) \to \mu(B)$ as $n \to \infty$ for any clopen set $B \subset X$. Because clopen sets are finite unions of cylinder sets, it suffices to show that if $D \subset \Gamma$ is a finite subset and $\xi \in \{0,1\}^D$ then

$\lim_{n\to\infty} \mu_n^\chi([\xi]) = \mu([\xi])$ where $[\xi]$ is the cylinder set $\{x \in X : x \restriction D = \xi\}$. We can further assume D to be a connected finite union of hyperedges with $1_\Gamma \in D$. The proof now follows from Lemma 3.6.3 since

$$\mu_n^\chi([\xi]) = n^{-1} \sum_{v\in V} \mathbb{P}_n^\chi\left(F_{D,\xi,v}\right) = n^{-1} \sum_{v\in\chi^{-1}(\xi(1_\Gamma))} \mathbb{P}_n^\chi\left(F_{D,\xi,v}\right) = (1/2)\mathbb{P}_n^\chi\left(F_{D,\xi,v_0}\right)$$

where, in the last equality, v_0 is any vertex in $\chi^{-1}(\xi(1_\Gamma))$. These equations are justified as follows. By symmetry, if $v_0, v_1 \in \chi^{-1}(\xi(1_\Gamma))$, then $\mathbb{P}_n^\chi\left(F_{D,\xi,v_0}\right) = \mathbb{P}_n^\chi\left(F_{D,\xi,v_1}\right)$. On the other hand, if $v \notin \chi^{-1}(\xi(1_\Gamma))$, then $\mathbb{P}_n^\chi\left(F_{D,\xi,v}\right) = 0$. Lastly, $|\chi^{-1}(\xi(1_\Gamma))| = n/2$.

□

3.6.2 The density of the rigid set

This subsection proves Lemma 3.5.1. So we assume the hypotheses of Proposition 3.4.11. An element $x \in X$ is a 2-coloring of the Cayley hyper-tree of Γ. Interpreted as such, $C_l(x), A_l(x), A'_l(x)$ are well-defined subsets of Γ (see §3.5 to recall the definitions).

For $l \in \mathbb{N} \cup \{\infty\}$, let

$$\begin{aligned}
\tilde{C}_l &= \{x \in X :\ 1_\Gamma \in C_l(x)\},\\
\tilde{A}_l &= \{x \in X :\ 1_\Gamma \in A_l(x)\},\\
\tilde{A}'_l &= \{x \in X :\ 1_\Gamma \in A'_l(x)\}.
\end{aligned}$$

Recall that $\lambda_0 = \frac{1}{2^{k-1}-1}$ and $\lambda = d\lambda_0$. Since we assume the hypothesis of Prop. 3.4.11, λ is asymptotic to $\log(2)k$ as $k \to \infty$.

Proposition 3.6.6.

$$\begin{aligned}\mu(\tilde{C}_\infty) &\geq 1 - \lambda^2 e^{-\lambda} + O(k^6 2^{-2k}),\\ \mu(\tilde{C}_\infty \cup \tilde{A}_\infty) &\geq 1 - e^{-\lambda} + O(k^4 2^{-2k}).\end{aligned}$$

Proof. For brevity, let $e_i \subset \Gamma$ be the subgroup generated by s_i. So e_i is a hyper-edge of the Cayley hyper-tree. Let $F_l^i \subset X$ be the set of all x such that

1. 1_Γ supports the edge e_i with respect to x and
2. $e_i \setminus \{1_\Gamma\} \subset C_l(x)$.

Since $C_{l+1}(x) \subset C_l(x)$, it follows that $F_{l+1}^i \subset F_l^i$. The events F_l^i for $i = 1, \ldots, d$ are i.i.d. Let $p_l = \mu(F_l^i)$ be their common probability.

We write $\text{Prob}(\text{Bin}(n,p) = m) = \binom{n}{m} p^m (1-p)^{n-m}$ for the probability that a binomial random variable with n trials and success probability p equals m. Since the events $F_{l-1}^1, \ldots, F_{l-1}^d$ are i.i.d., $\tilde{A}_l$ is the event that either 1 or 2 of these events occur and $\tilde{C}_l$ is the event that at least 3 of these events occur, it follows that

$$\mu(\tilde{A}_l) = \text{Prob}(\text{Bin}(d, p_{l-1}) \in \{1, 2\}).$$

$$\mu(\tilde{C}_l) = \text{Prob}(\text{Bin}(d, p_{l-1}) \geq 3).$$

Thus

$$\mu(\tilde{C}_l \cup \tilde{A}_l) = \text{Prob}(\text{Bin}(d, p_{l-1}) > 0).$$

Claim 3.6.7. *$p_0 = \lambda_0$ and for $l \geq 0$, $p_{l+1} = f(p_l)$ where*

$$f(t) = \lambda_0 \,\text{Prob}(\text{Bin}(d-1, t) \geq 3)^{k-1}.$$

Proof. To reduce notational clutter, let $F_l = F_l^1$. Note that $p_0 = \mu(F_0) = \lambda_0$ is the

probability that the edge e_1 is critical. So

$$p_{l+1} = \mu(F_0)\mu(F_{l+1} \mid F_0) = \lambda_0 \mu(F_{l+1} \mid F_0).$$

Conditioned on F_0, F_{l+1} is the event that $e_1 \setminus \{1_\Gamma\} \in C_{l+1}(x)$. By symmetry and the Markov property $\mu(F_{l+1} \mid F_0)$ is the $(k-1)$-st power of the probability that $s_1 \in C_{l+1}(x)$ given that 1_Γ supports e_1. By translation invariance, that probability is the same as the probability that $1_\Gamma \in C_{l+1}(x)$ given that 1_Γ does not support the edge e_1. By definition of $C_{l+1}(x)$ and the Markov property, this is the same as the probability that a binomial random variable with $(d-1)$ trials and success probability p_l is at least 3. This implies the claim. □

The next step is to bound $\text{Prob}(\text{Bin}(d-1,t) \geq 3)$ from below:

$$\begin{aligned}\text{Prob}(\text{Bin}(d-1,t) \geq 3) &= 1 - (1-t)^{d-1} - (d-1)t(1-t)^{d-2} - \binom{d-1}{2} t^2 (1-t)^{d-3} \\ &\geq 1 - e^{-(d-1)t}\left(1 + \frac{(d-1)t}{1-t} + \frac{(d-1)^2 t^2}{2(1-t)^2}\right). \qquad (3.37)\end{aligned}$$

The last inequality follows from the fact that $(1-t)^{d-1} \leq e^{-(d-1)t}$. This motivates the next claim:

Claim 3.6.8. *Suppose t is a number satisfying $\lambda_0\left(1-\lambda^2 e^{1-\lambda}\right)^{k-1} \leq t \leq \lambda_0$. Then for all sufficiently large k,*

$$\begin{aligned} 0 \leq \lambda - (d-1)t &\leq 1 \\ 1 + \frac{(d-1)t}{1-t} + \frac{(d-1)^2 t^2}{2(1-t)^2} &\leq (d-1)^2 t^2. \end{aligned}$$

Proof. The first inequality follows from:

$$\lambda - (d-1)t \geq \lambda - (d-1)\lambda_0 = \lambda_0 > 0.$$

The second inequality follows from:

$$\begin{aligned}\lambda - (d-1)t &\leq \lambda - (d-1)\lambda_0\left(1-\lambda^2 e^{1-\lambda}\right)^{k-1} \\ &= d\lambda_0 - (d-1)\lambda_0\left(1-\lambda^2 e^{1-\lambda}\right)^{k-1} \\ &\leq d\lambda_0 - (d-1)\lambda_0(1-(k-1)\lambda^2 e^{1-\lambda}) \\ &\leq \lambda_0 + (d-1)\lambda_0(k-1)\lambda^2 e^{1-\lambda} \leq \lambda_0 + k\lambda^3 e^{1-\lambda} \to_{k\to\infty} 0.\end{aligned}$$

The third line follows from the general inequality $(1-x)^{k-1} \geq 1-(k-1)x$ valid for all $x \in [0,1]$. To see the limit, observe that under the hypotheses of Proposition 3.4.11, $d \sim (\log(2)/2)k2^k$. So $\lambda \sim \log(2)k$. In particular, $k\lambda^3 e^{1-\lambda} \to 0$ and $\lambda_0 \to 0$ as $k \to \infty$. The implies the limit. Thus if k is large enough then the second inequality holds.

To see the last inequality, observe that since $t \leq \lambda_0$, $t \to 0$ as $k \to \infty$. On the other hand, $(d-1)t \sim \lambda \sim \log(2)k$. Thus $\frac{(d-1)t}{1-t}$ and $(d-1)t$ are asymptotic to $\log(2)k$. Since $1+\log(2)k+\frac{\log(2)^2k^2}{2} \leq \log(2)^2k^2$ for all sufficiently large k, this proves the last inequality assuming k is sufficiently large. □

Now suppose that t is as in Claim 3. Then

$$\begin{aligned}f(t) &\geq \lambda_0\left(1-e^{-(d-1)t}\left(1+\frac{(d-1)t}{1-t}+\frac{(d-1)^2t^2}{2(1-t)^2}\right)\right)^{k-1} \\ &\geq \lambda_0\left(1-e^{1-\lambda}(d-1)^2t^2\right)^{k-1} \geq \lambda_0\left(1-\lambda^2 e^{1-\lambda}\right)^{k-1}.\end{aligned}$$

The first inequality is implied by (3.37). The second and third inequalities follow from Claim 3. For example, since $\lambda-(d-1)t \leq 1$, $e^{-(d-1)t} \leq e^{1-\lambda}$.

Therefore, if p_l satisfies the bounds $\lambda_0\left(1-\lambda^2 e^{1-\lambda}\right)^{k-1} \leq p_l \leq \lambda_0$ then $f(p_l) = p_{l+1}$ satisfies the same bounds. Since $p_\infty = \lim_{l\to\infty} f^l(\lambda_0)$, it follows that

$$\lambda_0 \geq p_\infty \geq \lambda_0\left(1-\lambda^2 e^{1-\lambda}\right)^{k-1} = \lambda_0 + O(k^3 2^{-2k}). \tag{3.38}$$

Because $(1-\frac{t}{n})^n \leq e^{-t}$ for any $t, n > 0$,

$$\begin{aligned}\mu(\tilde{C}_\infty \cup \tilde{A}_\infty) &= \lim_{l\to\infty} \mu(\tilde{C}_l \cup \tilde{A}_l) = \lim_{l\to\infty} \text{Prob}(\text{Bin}(d, p_{l-1}) > 0) = \text{Prob}(\text{Bin}(d, p_\infty) > 0) \\ &= 1 - (1-p_\infty)^d \geq 1 - \exp(-p_\infty d) = 1 - e^{-\lambda} + O(k^4 2^{-2k}).\end{aligned}$$

The first equality occurs because $\tilde{C}_l \cup \tilde{A}_l$ decreases to $\tilde{C}_\infty \cup \tilde{A}_\infty$. By (3.37) and Claim 3.6.8 (with d in place of $d-1$),

$$\begin{aligned}\mu(\tilde{C}_\infty) &= \text{Prob}(\text{Bin}(d, p_\infty) \geq 3) \geq \text{Prob}(\text{Bin}(d, \lambda_0 + O(k^3 2^{-2k})) \geq 3) \\ &\geq 1 - \exp(-\lambda_0 d)\left(1 + \frac{d\lambda_0}{1-\lambda_0} + \frac{d^2\lambda_0^2}{2(1-\lambda_0)^2}\right) + O(k^6 2^{-2k}) \\ &\geq 1 - \lambda^2 e^{-\lambda} + O(k^6 2^{-2k}).\end{aligned}$$

□

Lemma 3.6.9. *$\mu(\tilde{A}'_\infty) = o(e^{-\lambda})$ where the implied limit is as $k \to \infty$ and η is bounded.*

Proof. As in the previous proof, let $e_i \subset \Gamma$ be the subgroup generated by s_i. So e_i is a hyper-edge of the Cayley hyper-tree.

Let $x \in X$. We say that an edge e is **attaching** (for x) if it is supported by a vertex $v \in A_\infty(x)$ and $e \setminus \{v\} \subset C_\infty(x)$. Let $F(x) = 0$ if $1_\Gamma \notin C_\infty(x)$. Otherwise, let $F(x)$ be the number of attaching edges containing 1_Γ. Then by translation invariance,

$$\mu(\tilde{A}'_\infty) \leq \sum_{m=2}^{d} m\mu(F(x) = m). \tag{3.39}$$

Let $G \subset X$ be the set of all x such that

1. e_1 is a critical edge supported by some vertex $v \neq 1_\Gamma$,

2. $e_1 \setminus \{v, 1_\Gamma\} \subset C_\infty(x)$,

3. $v \in A_\infty(x)$.

By the Markov property and symmetry,

$$\mu(F(x) = m) \leq \binom{d}{m} \mu(G)^m (1 - \mu(G))^{d-m}. \tag{3.40}$$

Let

- $G_1 \subset X$ be the set of all x such that e_1 is supported by s_1,
- $G_2 \subset X$ be the set of all x such that $e_1 \setminus \{s_1, 1_\Gamma\} \subset C_\infty(x)$,
- $G_3 \subset X$ be the set of all x such that $s_1 \in A_\infty(x)$.

By symmetry

$$\mu(G) = (k-1)\mu(G_3|G_2 \cap G_1)\mu(G_2|G_1)\mu(G_1).$$

Conditioned on $G_1 \cap G_2$, if G_3 occurs then there are no more than 1 attaching edge e supported by s_1 with $e \neq e_1$. By the Markov property and symmetry,

$$\mu(G_3|G_2 \cap G_1) \leq \text{Prob}(\text{Bin}(d-1, p_\infty) \leq 1) = O(\lambda e^{-\lambda})$$

where we have used (3.38). Also $\mu(G_1) = \lambda_0$. Thus $\mu(G) \leq O(k^2 e^{-2\lambda})$. So (3.39) and (3.40) along with straightforward estimates imply $\mu(\tilde{A}'_\infty) = o(e^{-\lambda})$.

□

Lemma 3.6.10. $\limsup_{l\to\infty} \mu(\tilde{A}'_l) \leq \mu(\tilde{A}'_\infty)$.

Proof. Given a coloring $\chi : \Gamma \to \{0, 1\}$ of the Cayley hyper-tree and $l \in \mathbb{N}$, define $A''_l(\chi) = \cup_{m \geq l} A'_m(\chi)$. Also define $\tilde{A}''_l = \{x \in X : \ 1_\Gamma \in A''_l(x)\}$. Since $\tilde{A}''_l \supset \tilde{A}'_l$ and the sets $\tilde{A}''_l$ are decreasing in l, it suffices to prove that $\cap_{l \geq 0} \tilde{A}''_l \subset \tilde{A}'_\infty$.

Suppose $x \in \cap_{l \geq 0} \tilde{A}''_l$. Then there exists an infinite set $S \subset \mathbb{N}$ such that $x \in \tilde{A}'_l$ $(\forall l \in S)$. So $1_\Gamma \in A'_l(x)$ $(\forall l \in S)$. So for each $l \in S$, there exist $g_l \in A_l(x) \setminus \{1_\Gamma\}$ and hyper-edges $e_l, f_l \subset \Gamma$ such that

1. 1_Γ supports e_l (with respect to x),

2. g_l supports f_l (with respect to x),

3. $e_l \cup f_l \setminus \{1_\Gamma, g_l\} \subset C_{l-1}(x)$,

4. $e_l \cap f_l \neq \emptyset$.

Because $e_l \cap f_l \neq \emptyset$, g_l is necessarily contained in the finite set $\{s_i^{p_i} s_j^{p_j} : 1 \leq i, j \leq d, 0 \leq p_i \leq k\}$. So after passing to an infinite subset of S if necessary, we may assume there is a fixed element $g \in \Gamma$ such that $g = g_l$ $(\forall l \in S)$. Similarly, we may assume there are edges $e, f \subset \Gamma$ such that $e_l = e$ and $f_l = f$ $(\forall l \in S)$.

Observe that $1_\Gamma \notin C_\infty(x)$ because $1_\Gamma \in A_l(x)$ implies $1_\Gamma \notin C_l(x)$ $(\forall l \in S)$. Similarly, $g \notin C_\infty(x)$. Because $e_l \cup f_l \setminus \{1_\Gamma, g_l\} \subset C_l(x)$ $(\forall l \in S)$ and the sets $C_l(x)$ are decreasing in l, it follows that $e \cup f \setminus \{1_\Gamma, g\} \subset C_\infty(x)$. Therefore $\{1_\Gamma, g\} \subset A_\infty(x)$. This verifies all of the conditions showing that $1_\Gamma \in A'_\infty(x)$ and therefore $x \in \tilde{A}'_\infty$ as required.

□

We can now prove Lemma 3.5.1:

Proof of Lemma 3.5.1. Observe that the sets $\tilde{C}_l, \tilde{A}_l, \tilde{A}'_l$ are clopen for finite l. By Lemma 3.6.1,

$$\lim_{\delta \searrow 0} \liminf_{n \to \infty} \mathbb{P}_n^\chi \left(\left| \frac{|C_l(\chi) \cup A_l(\chi) \setminus A'_l(\chi)|}{n} - \mu\left(\tilde{C}_l \cup \tilde{A}_l \setminus \tilde{A}'_l\right) \right| < \delta \right) = 1. \quad (3.41)$$

for any finite l. Since $\tilde{A}_\infty \cup \tilde{C}_\infty$ is the decreasing limit of $\tilde{A}_l \cup \tilde{C}_l$, Lemma 3.6.10 implies

$$\liminf_{l \to \infty} \mu(\tilde{C}_l \cup \tilde{A}_l \setminus \tilde{A}'_l) \geq \mu(\tilde{C}_\infty \cup \tilde{A}_\infty \setminus \tilde{A}'_\infty).$$

By Proposition 3.6.6 and Lemma 3.6.9,

$$\mu(\tilde{C}_\infty \cup \tilde{A}_\infty \setminus \tilde{A}'_\infty) \geq 1 - e^{-\lambda} + o(e^{-\lambda}).$$

Together with (3.41), this implies the lemma. □

3.7 Rigid vertices

This section proves Lemma 3.5.3. So we assume the hypotheses of Proposition 3.4.11.

As in the previous section, fix an equitable coloring $\chi : V \to \{0, 1\}$. We assume $|V| = n$ and let $\sigma : \Gamma \to \mathrm{Sym}(V)$ be a uniformly random uniform homomorphism conditioned on the event that χ is proper with respect to σ.

Lemma 3.7.1 (Expansivity Lemma)**.** *There is a constant $k_0 > 0$ such that the following holds. If $k \geq k_0$ then with high probability (with respect to the planted model), as $n \to \infty$, for any $T \subset V$ with $|T| \leq 2^{-k/2}n$ the following is true. For a vertex v let E_v denote the set of hyperedges supported by v. Let E_T be the set of all edges $e \in \cup_{v \in T} E_v$ such that $|e \cap T| \geq 2$. Then*

$$\#E_T \leq 2\#T.$$

Proof.

Claim 3.7.2. *There exists $k_0 \in \mathbb{N}$ such that $k \geq k_0$ implies*

- $k/2 \leq \lambda \leq k$,
- $1/2 - k2^{1-k/2} \geq \frac{1}{\sqrt{8}}$,
- *and for any $0 < t \leq 2^{-k/2}$ and $k/2 \leq \lambda' \leq k$*

$$H(t, 1-t) + \lambda' H(2t/\lambda', 1 - 2t/\lambda') + 2t\log(4k) + 4t\log(t) \leq 0.9t\log(t).$$

Proof. Recall that $\lambda = \log(2)k + O(k2^{-k})$. So the first two requirements are immediate for k_0 large enough.

We estimate each of the first three terms on the left as follows. Because $1 = \lim_{t \searrow 0} \frac{H(t,1-t)}{-t\log(t)}$, there exists $k_0 \in \mathbb{N}$ such that $k \geq k_0$ implies $\frac{H(t,1-t)}{-t\log(t)} \leq 1.01$.

Note,

$$\begin{aligned} \lambda' H(2t/\lambda', 1-2t/\lambda') &= -2t\log(2t/\lambda') - (\lambda' - 2t)\log(1-2t/\lambda') \\ &= -2t\log(2t/\lambda') + O(t) \leq -2t\log(t) + 2t\log(\lambda') + O(t) \\ &\leq -2t\log(t) + 2t\log(k) + O(t). \end{aligned}$$

So by making k_0 larger if necessary, we may assume

$$\frac{\lambda' H(2t/\lambda', 1-2t/\lambda')}{-t\log(t)} \leq 2.01.$$

Since

$$\frac{2t\log(4k)}{-t\log(t)} \leq \frac{2\log(4k)}{(k/2)\log(2)}$$

we may also assume $\frac{2t\log(4k)}{-t\log(t)} \leq 0.01$. Combining these inequalities, we obtain

$$\begin{aligned} & H(t,1-t) + \lambda' H(2t/\lambda', 1-2t/\lambda') + 2t\log(4k) + 4t\log(t) \\ \leq\ & (1.01 + 2.01 + 0.01 - 4)(-t\log(t)) \leq 0.9t\log(t). \end{aligned}$$

□

From now on, we assume $k \geq k_0$ with k_0 as above. To simplify notation, let $\zeta = 2^{-k/2}$. Given a $2d$-tuple $c = (c_{1,0}, c_{1,1} \ldots, c_{d,0}, c_{d,1})$ of natural numbers, let E_c be the event that there are exactly $c_{i,0}$ critical edges of the form $\{\sigma(s_i)^j(v) : 0 \leq j \leq k-1\}$ and supported by a vertex of color 0, and $c_{i,1}$ critical edges of the form $\{\sigma(s_i)^j(v) : 0 \leq j \leq k-1\}$ and supported by a vertex of color 1. We denote

$|c| = \sum_{1 \le i \le d, b \in \{0,1\}} c_{i,b}$. Let $\mathbb{P}^{\chi}_{c,n}$ be the planted model conditioned on E_c.

This measure can be constructed as follows. Let I_c be the set of triples (i, b, j) with $1 \le i \le d$, $b \in \{0, 1\}$ and $1 \le j \le c_{i,b}$. First choose edges $\{e_{i,b,j}\}_{(i,b,j) \in I_c}$ uniformly at random subject to the conditions:

1. each $e_{i,b,j} \subset [n]$ has cardinality k and $e_{i,b,j} \cap e_{i,b',j'} = \emptyset$ whenever $b \neq b'$ or $j \neq j'$,
2. each $e_{i,b,j}$ is critical and is supported by a vertex of color b with respect to χ.

Next choose a uniformly random uniform homomorphism σ subject to:

1. χ is a proper coloring with respect to σ,
2. each $e_{i,b,j}$ is of the form $\{\sigma(s_i)^j(v) : 0 \le j \le k-1\}$ with respect to σ,
3. the edges $\{e_{i,b,j}\}_{(i,b,j) \in I_c}$ are precisely the critical edges of χ with respect to σ.

Then σ is distributed according to $\mathbb{P}^{\chi}_{c,n}$.

For $1 \le l \le n$, let $\mathcal{T}_l$ be the collection of all subsets $T \subset V = [n]$ such that $|T| = l$ and $|E_T| > 2|T|$. Note that $\mathcal{T}_1$ is empty.

To prove the lemma, we claim it suffices to show the following:

Claim 3.7.3. *If n is sufficiently large and $kn/2 \le |c| \le kn$ then $\sum_{2 \le l \le \zeta n} \mathbb{E}^{\chi}_{c,n}[\#\mathcal{T}_l]$ tends to zero in n.*

We briefly return to the model $\mathbb{P}^{\chi}_n$ not conditioned on E_c. Let E be the set of all critical edges. By Lemma 3.6.1, with high probability in $\mathbb{P}^{\chi}_n$, $|E|$ is asymptotic to λn as $n \to \infty$. Let E'_n be the event that $(k/2)n \le |E| \le kn$. So $\mathbb{P}^{\chi}_n(E'_n) \to 1$ as $n \to \infty$.

By a first moment argument and the above paragraph, to prove the lemma it suffices to show that $\sum_{2 \le l \le \zeta n} \mathbb{E}^{\chi}_n(\#\mathcal{T}_l | E'_n)$ tends to 0 in n. Now $\mathbb{E}^{\chi}_n(\#\mathcal{T}_l | E'_n)$ is a convex combination of $\mathbb{E}^{\chi}_{c,n}(\#\mathcal{T}_l)$ over those c such that $kn/2 \le |c| \le kn$, so the lemma follows from Claim 3.7.3.

Before proving the above claim we need to prove another claim, which needs the following setup. For $s \in I_c$ and $T \subset V$, let $F_{T,s}$ be the event that e_s is supported by a vertex in T and $|e_s \cap T| \geq 2$. For $S \subset I_c$, let $F_{T,S} = \cap_{s \in S} F_{T,s}$. Note that for any $T \subset [n]$, the event $\{T \in \mathcal{T}_l\}$ is contained in $\cup_S F_{T,S}$ where the union is over all $S \subset I_c$ with $|S| = 2l$. So

$$\mathbb{E}^{\chi}_{c,n}[\#\mathcal{T}_l] \leq \sum_{S,T} \mathbb{P}^{\chi}_{c,n}(F_{T,S}). \tag{3.42}$$

where the sum is over all $T \subset [n]$ and $S \subset I_c$ with $|T| = l$ and $|S| = 2l$.

Before proving the claim above, we need to prove:

Claim 3.7.4. *For $1 \leq l \leq \zeta n$, any $T \subset V$ with cardinality $|T| = l$, any $S \subset I_c$ with $|S| \leq 2l - 1$, and any $s_0 \in I_c \setminus S$, one has $\mathbb{P}^{\chi}_{c,n}(F_{T,s_0} | F_{T,S}) \leq \frac{2kl^2}{n^2}$.*

Proof of Claim 3.7.4. For $s \in I_c$, let e_s be a random edge in $[n]$ with cardinality k as in the sampling algorithm above. Without loss of generality, we imagine that e_s for $s \in S$ has been chosen before e_{s_0}. Let $s_0 = (i_0, b_0, j_0)$. We say that an edge e is of type i if e is of the form $\{\sigma(s_i)^j(v) : 0 \leq j \leq k-1\}$. Let $S_0 \subset S$ be those edges of type i_0, and let $V_0 = \cup_{e \in S_0} e$,

- n_i be the number of vertices $v \in [n] \setminus V_0$ such that $\chi(v) = i$, and
- l_i be the number of vertices $v \in T \setminus V_0$ such that $\chi(v) = i$.

Let $e_0 = e_{s_0}$. The probability that e_0 is supported by a vertex v in T is l_{b_0}/n_{b_0} (this is conditioned on the edges e_s for $s \in S$).

Suppose first that e_0 is supported by a vertex v in T with $\chi(v) = 0$ (so $b_0 = 0$). Then the probability that $|e_0 \cap T| = 1$ is

$$\frac{\binom{n_1 - l_1}{k-1}}{\binom{n_1}{k-1}}.$$

It follows that for $b_0 = 0$

$$\mathbb{P}^{\chi}_{c,n}(F_{T,s_0}|F_{T,S}) = \frac{l_0}{n_0}\left(1 - \frac{\binom{n_1-l_1}{k-1}}{\binom{n_1}{k-1}}\right).$$

In order to bound this expression, consider

$$\begin{aligned}
\frac{\binom{n_1-l_1}{k-1}}{\binom{n_1}{k-1}} &= \left(\frac{n_1-l_1}{n_1}\right)\cdots\left(\frac{n_1-l_1-k+2}{n_1-k+2}\right) \\
&\geq \left(\frac{n_1-l_1-k+2}{n_1-k+2}\right)^{k-1} = \left(1-\frac{l_1}{n_1-k+2}\right)^{k-1} \\
&\geq 1 - \frac{(k-1)l_1}{n_1-k+2} \geq 1 - \frac{kl_1}{n_1}.
\end{aligned}$$

Thus,

$$\mathbb{P}^{\chi}_{c,n}(F_{T,s_0}|F_{T,S}) \leq \frac{kl_0l_1}{n_0n_1}.$$

A similar argument shows that the same bound above holds for the case $b_0 = 1$.

Because $l_0 + l_1 \leq |T| = l$, $(l_0+l_1)^2 - (l_0-l_1)^2 \leq l^2$, so that $l_0l_1 \leq l^2/4$. Note

$$n_0 \geq n/2 - 2kl \geq n(1/2 - k2^{1-k/2}) \geq n/\sqrt{8}$$

where we have used the assumption $l \leq \zeta n = 2^{-k/2}n$ and Claim 4. Similarly, $n_1 \geq n/\sqrt{8}$. Substitute these inequalities above to obtain

$$\mathbb{P}^{\chi}_{c,n}(F_{T,s_0}|F_{T,S}) \leq \frac{2kl^2}{n^2}.$$

□

We now prove Claim 5. Apply the chain rule and Claim 3.7.4 to obtain: if $S \subset I_c$

has $|S| = 2l$ and $T \subset [n]$ with $|T| = l$ then

$$\mathbb{P}^{\chi}_{c,n}(F_{T,S}) \leq \left(\frac{2kl^2}{n^2}\right)^{2l}.$$

By (3.42)

$$\mathbb{E}^{\chi}_{c,n}[\#\mathcal{T}_l] \leq \sum_{S,T} \mathbb{P}^{\chi}_{c,n}(F_{T,S}) \leq \binom{n}{l}\binom{|c|}{2l}\left(\frac{2kl^2}{n^2}\right)^{2l}$$

where the sum is over all $T \subset [n]$ and $S \subset I_c$ with $|T| = l$, $|S| = 2l$. Define t, λ' by $tn = l$ and $|c| = \lambda' n$. By hypothesis $k/2 \leq \lambda' \leq k$. Consider the following cases:
Case 1: $2 \leq l \leq n^{0.1}$. Then we make the following estimates:

$$\begin{aligned}\binom{n}{l} &\leq n^l,\\ \binom{|c|}{2l} &\leq (kn)^{2l}.\end{aligned}$$

It follows that $\mathbb{E}^{\chi}_{c,n}[\#\mathcal{T}_l] \leq \left(\frac{4k^4l^4}{n}\right)^l$, which is bounded by $n^{-1.1}$ for large enough n.
Case 2: $l > n^{0.1}$. We make the following estimates:

$$\begin{aligned}\binom{n}{l} &= \exp(nH(t, 1-t) + 0.5\log(n) + O(1)),\\ \binom{|c|}{2l} &\leq \exp(\lambda' nH(2t/\lambda', 1 - 2t/\lambda') + 0.5\log(kn) + O(1))\end{aligned}$$

so that

$$\mathbb{E}^{\chi}_{c,n}[\#\mathcal{T}_l] \leq Ckn\exp(n(H(t, 1-t) + \lambda' H(2t/\lambda', 1 - 2t/\lambda') + 2t\log(2k) + 4t\log(t)))$$

for some constant C. For n sufficiently large and $t \leq 2^{-k/2}$, this is bounded above by $Ckn\exp(0.9nt\log(t)) \leq Ckn2^{-0.45kn^{0.1}}$ by Claim 3.7.2 and the choice of k_0 and l. It follows that in this range of l, $\mathbb{E}^{\chi}_{c,n}[\#\mathcal{T}_l]$ decays super-polynomially.

This proves Claim 3.7.3 and finishes the lemma.

□

Lemma 3.7.5. *Let $\rho > 0$. Then there exists L such that $l > L$ implies $C_l(\chi) \subset [n]$ is ρ-rigid (with high probability in the planted model as $n \to \infty$).*

Proof. Without loss of generality, we may assume that $0 < \rho < \mu(\tilde{C}_\infty)$.

Observe that the sets $\tilde{C}_l$ are clopen for finite l. By Lemma 3.6.1,

$$\lim_{\eta \searrow 0} \liminf_{n\to\infty} \mathbb{P}_n^\chi \left(\left| \frac{|C_l(\chi)|}{n} - \mu\left(\tilde{C}_l\right) \right| < \eta \right) = 1.$$

Since the sets $C_l(\chi)$ are decreasing with l, this implies the existence of L such that $l > L$ implies

$$\liminf_{n\to\infty} \mathbb{P}_n^\chi \left(\left| \frac{|C_l(\chi)|}{n} - \frac{|C_{l+1}(\chi)|}{n} \right| < \rho/3 \right) = 1.$$

Choose $l > L$. Let $\psi : V \to \{0, 1\}$ be a σ-proper coloring. Let

$$T_l = \{v \in C_l(\chi) : \ \chi(v) \neq \psi(v)\}.$$

Define T_{l+1} similarly. Since $|C_l(\chi) \setminus C_{l+1}(\chi)| < \rho n/3$ (with high probability) and $T_l \setminus T_{l+1} \subset C_l(\chi) \setminus C_{l+1}(\chi)$, it follows that $|T_l \setminus T_{l+1}| < \rho n/3$ (with high probability).

For every $v \in T_{l+1}$, let $F_v \subset E_v$ be the subset of χ-critical edges e such that $e \subset C_l(\chi)$.

We claim that if $v \in T_{l+1}$ then $F_v \subset E_{T_l}$ where

$$E_{T_l} = \{e \in \cup_{v \in T_l} E_v : \ |e \cap T_l| \geq 2\}.$$

Since $v \in T_{l+1}$, $\psi(v) \neq \chi(v)$. If $e \in F_v$ then v supports e with respect to χ. Therefore because $\psi : [n] \to \{0, 1\}$ is a proper coloring, there must exist a vertex $w \in e \setminus \{v\}$ such that $\psi(w) \neq \psi(v)$. This, combined with $\chi(w) \neq \chi(v)$, implies $\psi(w) \neq \chi(w)$

since there are only two possible colors. Since $\{v, w\} \subset e \subset C_l(\chi)$, this means that $|e \cap T_l| \geq 2$ and therefore $e \in E_{T_l}$, which proves the claim.

For every $v \in T_{l+1}$, $|F_v| \geq 3$ by the definition of the sets $C_l(\chi)$. Since edges can only be supported by one vertex, the sets F_v are pairwise disjoint. So

$$|E_{T_l}| \geq \left| \bigcup_{v \in T_{l+1}} F_v \right| \geq 3|T_{l+1}| \geq 3|T_l| - \rho n.$$

If $|T_l| > \rho n$ then $|E_{T_l}| \geq 3|T_l| - \rho n > 2|T_l|$. So it follows from Lemma 3.7.1 that (with high probability), $|T_l| > 2^{-k/2}n$. Thus C_l is ρ-rigid. □

We can now prove Lemma 3.5.3.

Proof of Lemma 3.5.3. Let $\rho > 0$. By Lemma 3.7.5, there exists L such that $l > L$ implies $C_l(\chi)$ is $(\rho/3)$-rigid with high probability in the planted model as $n \to \infty$. So without loss of generality we condition on the event that $C_l(\chi)$ is $(\rho/3)$-rigid.

Now let $l - 1 > L$. Let $\psi : V \to \{0, 1\}$ be a σ-proper coloring. Let

$$T_{l-1} = \{v \in C_{l-1}(\chi) : \ \chi(v) \neq \psi(v)\}.$$

$$T_l = \{v \in C_l(\chi) : \ \chi(v) \neq \psi(v)\}.$$

$$T' = \{v \in A_l(\chi) \setminus A'_l(\chi) : \ \chi(v) \neq \psi(v)\}.$$

We claim that $|T_{l-1}| \geq |T'|$. To see this, let $v \in T'$. Then there exists an edge e supported by v (with respect to χ) with $e \setminus \{v\} \subset C_{l-1}(\chi)$. Since ψ is proper and $\psi(v) \neq \chi(v)$, there must exist a vertex $w \in e \setminus \{v\}$ with $\psi(w) \neq \chi(w)$. Necessarily, $w \in T_{l-1}$. So there exists a function $f : T' \to T_{l-1}$ such that $f(v)$ is contained in an edge e supported by v with $e \setminus \{v\} \subset C_{l-1}(\chi)$. Because $v \notin A'_l(\chi)$, f is injective. This proves the claim.

Now suppose that $|T_l \cup T'| > \rho n$. Since T_l and T' are disjoint, either $|T_l| > (\rho/3)n$

or $|T'| > (2\rho/3)n$. If $|T_l| > (\rho/3)n$, we are done because by assumption that $C_l(\chi)$ is $(\rho/3)$-rigid, $|T_l| > 2^{-k/2}n$ and so

$$|T_l \cup T'| > 2^{-k/2}n.$$

If $|T'| > (2\rho/3)n$ the claim implies $|T_{l-1}| > (2\rho/3)n$. Now in the proof of Lemma 3.7.5 we have shown that $|T_{l-1} \setminus T_l| < (\rho/3)n$, so $|T_l| > (\rho/3)n$ and we are again done. This proves the lemma. □

Appendices

Appendix A

Topological sofic entropy notions

In this appendix, we recall the notion of topological sofic entropy from [31] and prove that it coincides with the definition given in §3.1.

Let T be an action of Γ on a compact metrizable space X. So for $g \in \Gamma$, $T^g : X \to X$ is a homeomorphism and $T^{gh} = T^g T^h$. We will also denote this action by $\Gamma \curvearrowright X$. Let $\sigma : \Gamma \to \mathrm{Sym}(n)$ be a map, ρ be a pseudo-metric on X, $F \Subset \Gamma$ be finite and $\delta > 0$. For $x, y \in X^n$, let

$$\rho_\infty(x,y) = \max_i \rho(x_i, y_i), \quad \rho_2(x,y) = \left(\frac{1}{n} \sum_i \rho(x_i, y_i)^2 \right)^{1/2}$$

be pseudo-metrics on X^n. Also let

$$\mathrm{Map}(T, \rho, F, \delta, \sigma) = \{x \in X^n : \forall f \in F,\ \rho_2(T^f x, x \circ \sigma(f)) < \delta\}.$$

Informally, elements of $\mathrm{Map}(T, \rho, F, \delta, \sigma)$ are "good models" that approximate partial periodic orbits with respect to the chosen sofic approximation.

For a pseudo-metric space (Y, ρ), a subset $S \subset Y$ is (ρ, ϵ)-**separated** if for all $s_1 \neq s_2 \in S$, $\rho(s_1, s_2) \geq \epsilon$. Let $N_\epsilon(Y, \rho) = \max\{|S| : S \subset Y, S \text{ is } (\rho, \epsilon)\text{-separated}\}$ be the maximum cardinality over all (ρ, ϵ)-separated subsets of Y.

Given a sofic approximation Σ to Γ, we define

$$\tilde{h}_\Sigma(\Gamma \curvearrowright X, \rho) = \sup_{\epsilon>0} \inf_{F \Subset \Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log(N_\epsilon(\mathrm{Map}(T, \rho, F, \delta, \sigma_i), \rho_\infty))$$

where the symbol $F \Subset \Gamma$ means that F varies over all finite subsets of Γ.

We say that a pseudo-metric ρ on X is **generating** if for every $x \neq y$ there exists $g \in \Gamma$ such that $\rho(gx, gy) > 0$. By [31, Proposition 2.4], if ρ is continuous and generating, $\tilde{h}_\Sigma(T, \rho)$ is invariant under topological conjugacy and does not depend on the choice of ρ. So we define $\tilde{h}_\Sigma(T) = \tilde{h}_\Sigma(T, \rho)$ where ρ is any continuous generating pseudo-metric. The authors of [31] define the topological sofic entropy of $\Gamma \curvearrowright X$ to be $\tilde{h}_\Sigma(T)$. The main result of this appendix is:

Proposition A.0.1. *Let $\mathcal{A}$ be a finite set and $X \subset \mathcal{A}^\Gamma$ a closed shift-invariant subspace. Let T be the shift action of Γ on X. Then $h_\Sigma(\Gamma \curvearrowright X) = \tilde{h}_\Sigma(T)$ where $h_\Sigma(\Gamma \curvearrowright X)$ is as defined in §3.1.*

Proof. To begin, we choose a pseudo-metric on $\mathcal{A}^\Gamma$ as follows. For $x, y \in \mathcal{A}^\Gamma$, let $\rho(x, y) = 1_{x_{1_\Gamma} \neq y_{1_\Gamma}}$. Then ρ is continuous and generating. So $\tilde{h}_\Sigma(\Gamma \curvearrowright X) = \tilde{h}_\Sigma(\Gamma \curvearrowright X, \rho)$.

Let $\epsilon > 0$, $\mathcal{O} \subset \mathcal{A}^\Gamma$ be an open set. We first analyze $\Omega(\mathcal{O}, \epsilon, \sigma)$ from the definition of $h_\Sigma(\Gamma \curvearrowright X)$. Note that the topology for $\mathcal{A}^\Gamma$ is generated by the base $\mathcal{B} = \{[a] : a \in \mathcal{A}^F, F \Subset \Gamma\}$ where if $a \in \mathcal{A}^F$ then $[a] = \{x \in \mathcal{A}^\Gamma : x|_F = a\}$. In other words, open sets of $\mathcal{B}$ are those that specify a configuration on a finite subset of coordinates. For $F \Subset \Gamma$ let $\mathcal{O}(F) = \{y \in \mathcal{A}^\Gamma : \exists x \in X, y|_F = x|_F\} = \cup_{\substack{a \in \mathcal{A}^F \\ [a] \cap X \neq \emptyset}} [a]$ be the open set containing all elements containing some configuration that appears in X in the finite window F.

Claim A.0.2. *Every open superset $\mathcal{O} \supset X$ contains some open set of the form $\mathcal{O}(F)$.*

Because $\Omega(\mathcal{O}, \epsilon, \sigma)$ decreases as $\mathcal{O}$ decreases, it suffices to only consider open sets of the form $\mathcal{O}(F)$ in the definition of $h_\Sigma(\Gamma \curvearrowright X)$.

Proof. $\mathcal{O}$ is a union of elements in $\mathcal{B}$ and X is compact, so that there exists $X \subset \mathcal{O}' \subset \mathcal{O}$ with $\mathcal{O}'$ containing only finitely many base elements. Let F be the union of all coordinates specified by base elements in $\mathcal{O}'$. It follows that $\mathcal{O}'$ contains $\mathcal{O}(F)$. $\square$

Without loss of generality and for convenience we can assume that F is symmetric, i.e. $F = F^{-1}$, and contains the identity. This is because we can replace any F with the larger set $F \cup F^{-1} \cup \{1_\Gamma\}$, and both $\mathrm{Map}(T, \rho, F, \delta, \sigma_i)$ and $\Omega(\mathcal{O}(F), \epsilon, \sigma_i)$ are monotone decreasing in F.

Let $n = |V_i|$. We assume $\lim_{i\to\infty} |V_i| = \infty$. Now for each $x \in \Omega(\mathcal{O}(F), \epsilon, \sigma_i)$ we obtain an element $\tilde{x} \in X^n$ and then show that these partial orbits form a good estimate for $\tilde{h}_\Sigma$. Let $G(x) = \{v \in V_i : \Pi_v^{\sigma_i}(x) \in \mathcal{O}(F)\}$. For every $v \in G(x)$, choose some $\tilde{x}_v \in X$ that agrees with $\Pi_v^{\sigma_i}(x)$ on F. For $v \notin G(x)$ choose an arbitrary element $\tilde{x}_v \in X$. Thus $\tilde{x} \in X^n$.

Now for $v \in G(x), f \in F$, $T^f \tilde{x}_v(1_\Gamma) = \tilde{x}_v(f^{-1}) = x_{\sigma_i(f)v}$. On the other hand we also want $\tilde{x}_{\sigma_i(f)v}(1_\Gamma) = x_{\sigma_i(f)v}$, which is true if $v \in \sigma_i(f)^{-1}G(x)$ and $\sigma_i(1_\Gamma)\sigma_i(f)v = \sigma_i(f)v$. It follows that $\rho_2(T^f\tilde{x}, \tilde{x} \circ \sigma_i(f)) < \sqrt{2\epsilon}$.

Now consider separation of $\{\tilde{x} : x \in \Omega(\mathcal{O}(F), \epsilon, \sigma_i)\}$. We will show that a slightly smaller subset is $(\rho_\infty, 1)$-separated. By the pigeonhole principle there exists a subset $\bar{V}_i$ of size at least $(1-\epsilon)n$ such that $\Omega(\mathcal{O}(F), \epsilon, \sigma_i, \bar{V}_i) := \{x \in \Omega(\mathcal{O}(F), \epsilon, \sigma_i) : G(x) = \bar{V}_i\}$ has cardinality at least $e^{-n(H(\epsilon,1-\epsilon)+o(1))}\#\Omega(\mathcal{O}(F), \epsilon, \sigma_i)$. Furthermore, if $x, y \in \Omega(\mathcal{O}(F), \epsilon, \sigma_i, \bar{V}_i)$ then $\rho_\infty(\tilde{x}, \tilde{y}) = 1$ if $x(v) \neq y(v)$ for some $v \in \bar{V}_i \cap \mathrm{Fix}(1_\Gamma)$, where $\mathrm{Fix}(1_\Gamma) = \{v \in V_i : \sigma_i(1_\Gamma)v = v\}$. Since there are at most $|\mathcal{A}|^{(\epsilon+o(1))n}$ configurations in $\mathcal{A}^{V_i}$ with some fixed configuration on $\bar{V}_i \cap \mathrm{Fix}(1_\Gamma)$, there exists a $(\rho_\infty, 1)$-separated subset of $\Omega(\mathcal{O}(F), \epsilon, \sigma_i, \bar{V}_i)$ of size at least $|\mathcal{A}|^{-(\epsilon+o(1))n}\#\Omega(\mathcal{O}(F)), \epsilon, \sigma_i, \bar{V}_i)$. It follows that

$$N_1(\mathrm{Map}(T, \rho, F, \sqrt{2\epsilon}, \sigma_i), \rho_\infty) \geq |\mathcal{A}|^{-(\epsilon+o(1))n} e^{-n(H(\epsilon,1-\epsilon)+o(1))}\#\Omega(\mathcal{O}(F), \epsilon, \sigma_i).$$

On the other hand, suppose we have some $\tilde{x} \in \mathrm{Map}(T, \rho, F, \delta, \sigma_i)$. This means that for every $f \in F$, there exists a set $\tilde{V}_i(f)$ of size $> (1-\delta^2)n$ such that for $v \in \tilde{V}_i(f)$, $\tilde{x}_{\sigma_i(f)v}(1_\Gamma) = T^f \tilde{x}_v(1_\Gamma) = \tilde{x}_v(f^{-1})$. Let $\tilde{V}_i = \cap_{f \in F} \tilde{V}_i(f)$. Then $|\tilde{V}_i| > (1 - |F|\delta^2)n$ and for $v \in \tilde{V}_i$, for every $f \in F$, $\tilde{x}_{\sigma_i(f)v}(1_\Gamma) = T^f \tilde{x}_v(1_\Gamma) = \tilde{x}_v(f^{-1})$.

Define $x \in \mathcal{A}^{V_i}$ by $x_v = \tilde{x}_v(1_\Gamma)$. Then for any fixed $v \in \tilde{V}_i$, for every $f \in F$, $\Pi_v^{\sigma_i}(x)(f) = x_{\sigma_i(f^{-1})v} = \tilde{x}_{\sigma_i(f^{-1})v}(1_\Gamma) = T^{f^{-1}} \tilde{x}_v(1_\Gamma) = \tilde{x}_v(f)$. Since $\tilde{x}_v \in X$, it follows that $x \in \Omega(\mathcal{O}(F), \delta^2|F|, \sigma_i)$.

Also note that $\tilde{x}, \tilde{y} \in \mathrm{Map}(T, \rho, F, \delta, \sigma_i)$ are (ρ, ϵ)-separated for any $\epsilon \leq 1$ if and only if $\tilde{x}_v(1_\Gamma) \neq \tilde{y}_v(1_\Gamma)$ for some $v \in V_i$, so that $x \neq y$. It follows that

$$N_\epsilon(\mathrm{Map}(T, \rho, F, \delta, \sigma_i), \rho_\infty) \leq \#\Omega(\mathcal{O}(F), \delta^2|F|, \sigma_i).$$

Note that in the definitions of h_Σ and $\tilde{h}_\Sigma$, F is fixed with respect to δ.

□

Appendix B

Concentration for the planted model

Definition B.0.1 (Hamming metrics)**.** *Define the* **normalized Hamming metric** $d_{\mathrm{Sym}(n)}$ *on* $\mathrm{Sym}(n)$ *by*

$$d_{\mathrm{Sym}(n)}(\sigma_1, \sigma_2) = n^{-1}\#\{i \in [n] : \ \sigma_1(i) \neq \sigma_2(i)\}.$$

Define the **normalized Hamming metric** d_{Hom} *on* $\mathrm{Hom}(\Gamma, \mathrm{Sym}(n))$ *by*

$$d_{\mathrm{Hom}}(\sigma_1, \sigma_2) = \sum_{i=1}^{d} d_{\mathrm{Sym}(n)}(\sigma_1(s_i), \sigma_2(s_i)).$$

The purpose of this section is to prove:

Theorem B.0.2. *There exist constants* $c, \lambda > 0$ *(depending only on* k, d*) such that for every* $\delta > 0$ *there exists* N_δ *such that for all* $n > N_\delta$*, for every* 1*-Lipschitz* $f : \mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n)) \to \mathbb{R}$*,*

$$\mathbb{P}_n^\chi \left(|f - \mathbb{E}_n^\chi[f]| > \delta \right) \leq c \exp(-\lambda \delta^2 n).$$

B.0.1 General considerations

To begin the proof we first introduce some general-purpose tools.

Definition B.0.3. *A* **metric measure space** *is a triple* (X, d_X, μ) *where* (X, d_X) *is a metric space and* μ *is a Borel probability measure on* X*. We will say* (X, d_X, μ) *is* (c, λ)**-concentrated** *if for any 1-Lipschitz function* $f : X \to \mathbb{R}$*,*

$$\mu\left(\left|f - \int f \, d\mu\right| > \epsilon\right) < ce^{-\lambda\epsilon^2}.$$

If (X, d_X, μ) *is* (c, λ)*-concentrated and* $f : X \to \mathbb{R}$ *is* L*-Lipschitz, then since* f/L *is* 1*-Lipschitz,*

$$\mu\left(\left|f - \int f \, d\mu\right| > \epsilon\right) = \mu\left(\left|f/L - \int f/L \, d\mu\right| > \epsilon/L\right) < c\exp(-\lambda\epsilon^2/L^2). \quad \text{(B.1)}$$

Lemma B.0.4. *Let* (X, d_X, μ) *be* (c, λ)*-concentrated. If* $\phi : X \to Y$ *is an* L*-Lipschitz map onto a measure metric space* (Y, d_Y, ν) *and* $\nu = \phi_*\mu$ *is the push-forward measure, then* (Y, d_Y, ν) *is* $(c, \lambda/L^2)$*-concentrated.*

Proof. This follows from the observation that if $f : Y \to \mathbb{R}$ is 1-Lipschitz, then the pullback $f \circ \phi : X \to \mathbb{R}$ is L-Lipschitz. So equation (B.1) implies

$$\nu\left(\left|f - \int f \, d\nu\right| > \epsilon\right) = \mu\left(\left|f \circ \phi - \int f \circ \phi \, d\mu\right| > \epsilon\right) < c\exp(-\lambda\epsilon^2/L^2).$$

□

The next lemma is concerned with the following situation. Suppose $X = \sqcup_{i \in I} X_i$ is a finite disjoint union of spaces X_i. Even if we have good concentration bounds on the spaces X_i, this does not imply concentration on X because it is possible that a 1-Lipschitz function f will have different means when restricted to the X_i's. However, if most of the mass of X is concentrated on a sub-union $\cup_{j \in J} X_j$ (for some $J \subset I$)

and the sets X_j are all very close to each other, then there is a weak concentration inequality on X.

Lemma B.0.5. *Let (X, d_X, μ) be a measure metric space with diameter ≤ 1. Suppose $X = \sqcup_{i \in I} X_i$ is a finite disjoint union of spaces X_i, each with positive measure ($\mu(X_i) > 0$). Let μ_i be the induced probability measure on X_i. Suppose there exist $J \subset I$ and constants $\eta, \delta, \lambda, c > 0$ satisfying:*

1. *$\mu(\cup_{j \in J} X_j) \geq 1 - \eta \geq 1/2$.*
2. *For every $j, k \in J$, there exists a measure $\mu_{j,k}$ on $X_j \times X_k$ with marginals μ_j, μ_k respectively such that*

$$\mu_{j,k}(\{(x_j, x_k) : \ d_X(x_j, x_k) \leq \delta\}) = 1.$$

3. *For each $j \in J$, (X_j, d_X, μ_j) is (c, λ)-concentrated.*

Then for every 1-Lipschitz function $f : X \to \mathbb{R}$ and every $\epsilon > \delta + 2\eta$,

$$\mu\left(\left|f - \int f \ d\mu\right| > \epsilon\right) \leq \eta + c \exp\left(-\lambda \left(\epsilon - \delta - 2\eta\right)^2\right).$$

Proof. Let $f : X \to \mathbb{R}$ be a 1-Lipschitz function. After adding a constant to f if necessary, we may assume $\int f \ d\mu = 0$. Note that the mean of f is a convex combination of its restrictions to the X_i's:

$$\begin{aligned} 0 = \int f(x) \ d\mu(x) &= \sum_{i \in I} \mu(X_i) \int f(x_i) \ d\mu_i(x_i) \\ &= \sum_{i \in I \setminus J} \mu(X_i) \int f(x_i) \ d\mu_i(x_i) + \sum_{j \in J} \mu(X_j) \int f(x_j) d\mu_j(x_j). \end{aligned}$$

Since f is 1-Lipschitz with zero mean, $|f| \leq \mathrm{diam}(X) \leq 1$. So

$$\begin{aligned}\left|\mu(\cup_{j\in J}X_j)^{-1}\sum_{j\in J}\mu(X_j)\int f(x_j)\ d\mu_j(x_j)\right| &= \left|\mu(\cup_{j\in J}X_j)^{-1}\sum_{i\in I\setminus J}\mu(X_i)\int f(x_i)\ d\mu_i(x_i)\right| \\ &\leq \frac{\eta}{1-\eta} \leq 2\eta\end{aligned}$$

where the last inequality uses that $\mu(\cup_{j\in J}X_j) \geq 1-\eta$ and $\eta \leq 1/2$.

For any $j,k \in J$, the μ_j and μ_k-means of f are δ-close:

$$\begin{aligned}\left|\int f(x_j)\ d\mu_j(x_j) - \int f(x_k)\ d\mu_k(x_k)\right| &= \left|\int f(x_j) - f(x_k)\ d\mu_{j,k}(x_j,x_k)\right| \\ &\leq \int |f(x_j)-f(x_k)|\ d\mu_{j,k}(x_j,x_k) \leq \delta.\end{aligned}$$

So for any $j_0 \in J$,

$$\left|\int f(x_{j_0})\ d\mu_{j_0}(x_{j_0}) - \mu(\cup_{j\in J}X_j)^{-1}\sum_{j\in J}\mu(X_j)\int f(x_j)\ d\mu_j(x_j)\right| \leq \delta.$$

Combined with the previous estimate, this gives

$$\left|\int f(x_{j_0})\ d\mu_{j_0}(x_{j_0})\right| \leq \delta + 2\eta.$$

Now we estimate the μ-probability that f is $> \epsilon$ (assuming $\epsilon > \delta + 2\eta$):

$$\begin{aligned}
\mu\left(\left|f-\int f\ d\mu\right|>\epsilon\right) &= \mu(|f|>\epsilon)\\
&\leq \eta+\sum_{j\in J}\mu_j(|f|>\epsilon)\mu(X_j)\\
&\leq \eta+\sum_{j\in J}\mu_j\left(\left|f-\int f(x_j)\ d\mu_j(x_j)\right|>\epsilon-\left|\int f(x_j)\ d\mu_j(x_j)\right|\right)\mu(X_j)\\
&\leq \eta+\sum_{j\in J}\mu_j\left(\left|f-\int f(x_j)\ d\mu_j(x_j)\right|>\epsilon-\delta-2\eta\right)\mu(X_j)\\
&\leq \eta+c\exp\left(-\lambda\left(\epsilon-\delta-2\eta\right)^2\right).
\end{aligned}$$

□

The next lemma is essentially the same as [32, Proposition 1.11]. We include a proof for convenience.

Lemma B.0.6. *[32] Suppose (X,d_X,μ) is (c_1,λ_1)-concentrated and (Y,d_Y,ν) is (c_2,λ_2)-concentrated. Define a metric on $X\times Y$ by $d_{X\times Y}((x_1,y_1),(x_2,y_2))=d_X(x_1,x_2)+d_Y(y_1,y_2)$. Then $(X\times Y,d_{X\times Y},\mu\times\nu)$ is $(c_1+c_2,\min(\lambda_1,\lambda_2)/4)$-concentrated.*

Proof. Let $F:X\times Y\to\mathbb{R}$ be 1-Lipschitz. For $y\in Y$, define $F^y:X\to\mathbb{R}$ by $F^y(x)=F(x,y)$. Define $G:Y\to\mathbb{R}$ by $G(y)=\int F^y(x)\ d\mu(x)$. Then F^y and G are 1-Lipschitz.

If $|F(x,y)-\int F\ d\mu\times\nu|>\epsilon$ then either $|F^y(x)-\int F^y\ d\mu|>\epsilon/2$ or $|G(y)-\int G\ d\nu|>\epsilon/2$. Thus

$$\begin{aligned}
&\mu\times\nu\left(\left\{(x,y):\ \left|F(x,y)-\int F\ d\mu\times\nu\right|>\epsilon\right\}\right)\\
\leq\ &\mu\times\nu\left(\left\{(x,y):\ \left|F^y(x)-\int F^y\ d\mu\right|>\epsilon/2\right\}\right)+\nu\left(\left\{y:\ \left|G(y)-\int G\ d\nu\right|>\epsilon/2\right\}\right)\\
\leq\ &c_1e^{-\lambda_1\epsilon^2/4}+c_2e^{-\lambda_2\epsilon^2/4}\leq(c_1+c_2)\exp(-\min(\lambda_1,\lambda_2)\epsilon^2/4).
\end{aligned}$$

□

Lemma B.0.7. *Let (X, d_X, μ) and (Y, d_Y, ν) be metric-measure spaces. Suppose*

1. *X, Y are finite sets, μ and ν are uniform probability measures,*
2. *there is a surjective map $\Phi : X \to Y$ and a constant $C > 0$ such that $|\Phi^{-1}(y)| = C$ for all $y \in Y$,*
3. *(Y, d_Y, ν) is (c_1, λ_1)-concentrated,*
4. *for each $y \in Y$, the fiber $\Phi^{-1}(y)$ is (c_2, λ_2)-concentrated (with respect to the uniform measure on $\Phi^{-1}(y)$ and the restricted metric),*
5. *for each $y_1, y_2 \in Y$ there is a probability measure μ_{y_1,y_2} on $\Phi^{-1}(y_1) \times \Phi^{-1}(y_2)$ with marginals equal to the uniform measures on $\Phi^{-1}(y_1)$ and $\Phi^{-1}(y_2)$ such that*

$$\mu_{y_1,y_2}(\{(x_1, x_2) : \ d_X(x_1, x_2) \leq d_Y(y_1, y_2)\}) = 1.$$

Then (X, d_X, μ) is $(c_1 + c_2, \min(\lambda_1, \lambda_2)/4)$-concentrated.

Proof. Let $f : X \to \mathbb{R}$ be 1-Lipschitz. Let $\mathbb{E}[f|Y] : Y \to \mathbb{R}$ be its conditional expectation defined by

$$\mathbb{E}[f|Y](y) = |\Phi^{-1}(y)|^{-1} \sum_{x \in \Phi^{-1}(y)} f(x).$$

Also let $\mathbb{E}[f] = |X|^{-1} \sum_{x \in X} f(x)$ be its expectation.

We claim that $\mathbb{E}[f|Y]$ is 1-Lipschitz. So let $y_1, y_2 \in Y$. By hypothesis (5)

$$\begin{aligned} \mathbb{E}[f|Y](y_1) - \mathbb{E}[f|Y](y_2) &= \int f(x_1) - f(x_2) \ d\mu_{y_1,y_2}(x_1, x_2) \\ &\leq \int d_X(x_1, x_2) \ d\mu_{y_1,y_2}(x_1, x_2) \\ &\leq d_Y(y_1, y_2). \end{aligned}$$

The first inequality holds because f is 1-Lipschitz and the second by hypothesis (5). This proves $\mathbb{E}[f|Y]$ is 1-Lipschitz.

Let $\epsilon > 0$. Because Φ is C-to-1, it pushes forward the measure μ to ν. Because (Y, d_Y, ν) is (c_1, λ_1)-concentrated,

$$\mu\left(\left|\mathbb{E}[f|Y] \circ \Phi - \int f \, d\mu\right| > \epsilon/2\right) = \nu\left(|\mathbb{E}[f|Y] - \mathbb{E}[f]| > \epsilon/2\right) < c_1 e^{-\lambda_1 \epsilon^2/4}. \quad \text{(B.2)}$$

Because each fiber $\Phi^{-1}(y)$ is (c_2, λ_2)-concentrated, for any $y \in Y$,

$$|\Phi^{-1}(y)|^{-1} \# \left\{x \in \Phi^{-1}(y) : \ |f(x) - \mathbb{E}[f|Y](y)| > \epsilon/2\right\} < c_2 e^{-\lambda_2 \epsilon^2/4}.$$

Average this over $y \in Y$ to obtain

$$\mu\left(\{x \in X : \ |f(x) - \mathbb{E}[f|Y](\Phi(x))| > \epsilon/2\right) < c_2 e^{-\lambda_2 \epsilon^2/4}.$$

Combine this with (B.2) to obtain

$$\begin{aligned} \mu\left(\left|f - \int f \, d\mu\right| > \epsilon\right) &\leq \mu\left(|f - \mathbb{E}[f|Y](\Phi(x))| > \epsilon/2\right) + \mu\left(\left|\mathbb{E}[f|Y](\Phi(x)) - \int f \, d\mu\right| > \epsilon/2\right) \\ &\leq c_2 e^{-\lambda_2 \epsilon^2/4} + c_1 e^{-\lambda_1 \epsilon^2/4} \end{aligned}$$

which implies the lemma.

□

B.0.2 Specific considerations

Given an equitable coloring $\chi : [n] \to \{0, 1\}$, let H_χ be the stabilizer of χ:

$$H_\chi = \{g \in \mathrm{Sym}(n) : \ \chi(gv) = \chi(v) \ \forall v \in [n]\}.$$

Lemma B.0.8. *The group H_χ is $(4, n/16)$-concentrated (when equipped with the uniform probability measure and the restriction of the normalized Hamming metric $d_{\mathrm{Sym}(n)}$).*

Proof. The group H_χ is isomorphic to the direct product $\mathrm{Sym}(\chi^{-1}(0)) \times \mathrm{Sym}(\chi^{-1}(1))$ which is isomorphic to $\mathrm{Sym}(n/2)^2$. By [32, Corollary 4.3], $\mathrm{Sym}(n/2)$ is $(2, n/16)$-concentrated. So the result follows from Lemmas B.0.6 and B.0.4. This uses that the inclusion map from $\mathrm{Sym}(n/2)^2$ to itself is $(1/2)$-Lipschitz when the source is equipped with the sum of the $d_{\mathrm{Sym}(n/2)}$-metrics and the target equipped with the $d_{\mathrm{Sym}(n)}$ metric. □

We need to show that certain subsets of the group $\mathrm{Sym}(n)$ are concentrated. To define these subsets, we need the following terminology.

Recall that a **k-partition** of $[n]$ is an unordered partition $\pi = \{P_1, \ldots, P_{n/k}\}$ of $[n]$ such that each P_i has cardinality k. Let $\mathrm{Part}(n, k)$ be the set of all k-partitions of $[n]$. The group $\mathrm{Sym}(n)$ acts on $\mathrm{Part}(n, k)$ by $g\pi = \{gP_1, \ldots, gP_{n/k}\}$.

Let $\sigma \in \mathrm{Sym}(n)$. The **orbit-partition** of σ is the partition $\mathrm{Orb}(\sigma)$ of $[n]$ into orbits of σ. For example, for any $v \in [n]$ the element of $\mathrm{Orb}(\sigma)$ containing v is $\{\sigma^i v : \ i \in \mathbb{Z}\} \subset [n]$. Let $\mathrm{Sym}(n, k) \subset \mathrm{Sym}(n)$ be the set of all permutations $\sigma \in \mathrm{Sym}(n)$ such that the orbit-partition of σ is a k-partition.

Recall from §3.3.1 that a k-partition π **has type** $\vec{t} = (t_j)_{j=0}^k \in [0,1]^{k+1}$ **with respect to a coloring** χ if the number of partition elements P of π with $|P \cap \chi^{-1}(1)| = j$ is $t_j n$. We will also say that a permutation $\sigma \in \mathrm{Sym}(n, k)$ **has type** $\vec{t} = (t_j)_{j=0}^k \in [0,1]^{k+1}$ **with respect to a coloring** χ if its orbit-partition $\mathrm{Orb}(\sigma)$ has type $\vec{t}$ with respect to χ.

Let $\mathrm{Sym}(n, k; \chi, \vec{t})$ be the set of all permutations $\sigma \in \mathrm{Sym}(n, k)$ such that σ has type $\vec{t}$ with respect to χ.

Lemma B.0.9. *The subset $\mathrm{Sym}(n, k; \chi, \vec{t})$ is either empty or $(6, \lambda n)$-concentrated*

(when equipped with the normalized Hamming metric $d_{\mathrm{Sym}(n)}$ and the uniform probability measure) where $\lambda > 0$ is a constant depending only on k.

Proof. Let $\mathrm{Part}(n, k; \chi, \vec{t})$ be the set of all (unordered) k-partitions of $[n]$ with type $\vec{t}$ (with respect to χ). We will consider this set as a metric space in which the distance between partitions $\pi, \pi' \in \mathrm{Part}(n, k; \chi, \vec{t})$ is $d(\pi, \pi') = \frac{k|\pi \triangle \pi'|}{2n}$ where $\triangle$ denotes symmetric difference.

Let $\mathrm{Orb} : \mathrm{Sym}(n, k; \chi, \vec{t}) \to \mathrm{Part}(n, k; \chi, \vec{t})$ be the map which sends a permutation to its orbit-partition. We will verify the conditions of Lemma B.0.7 with $X = \mathrm{Sym}(n, k; \chi, \vec{t})$, $Y = \mathrm{Part}(n, k; \chi, \vec{t})$ and $\Phi = \mathrm{Orb}$. Condition (1) is immediate.

Observe that Orb is surjective and constant-to-1. In fact for any partition $\pi \in \mathrm{Part}(n, k; \chi, \vec{t})$, $|\mathrm{Orb}^{-1}(\pi)| = (k-1)!^{n/k}$ since an element $\sigma \in \mathrm{Orb}^{-1}(\pi)$ is obtained by choosing a k-cycle for every part of π. To be precise, if $\pi = \{P_1, \ldots, P_{n/k}\}$ then $\mathrm{Orb}^{-1}(\pi)$ is the set of all permutations σ of the form $\sigma = \prod_{i=1}^{n/k} \sigma_i$ where σ_i is a k-cycle with support in P_i. This verifies condition (2) of Lemma B.0.7.

Observe that H_χ acts transitively on $\mathrm{Part}(n, k; \chi, \vec{t})$. Fix $\pi \in \mathrm{Part}(n, k; \chi, \vec{t})$ and define a map $\phi : H_\chi \to \mathrm{Part}(n, k; \chi, \vec{t})$ by $\phi(h) = h\pi$. We claim that ϕ is $k^2/2$-Lipschitz. Indeed, if $h_1, h_2 \in H_\chi$ then

$$\begin{aligned} d(h_1\pi, h_2\pi) &= \frac{k|h_1\pi \triangle h_2\pi|}{2n} \\ &\leq \frac{k^2 \#\{p \in [n] : \ h_1(p) \neq h_2(p)\}}{2n} \\ &= \frac{k^2}{2} d_{\mathrm{Sym}(n)}(h_1, h_2). \end{aligned}$$

Because H_χ is $(4, n/16)$-concentrated by Lemma B.0.8, Lemma B.0.4 implies $\mathrm{Part}(n, k; \chi, \vec{t})$ is $(4, n/4k^4)$-concentrated. This verifies condition (3) of Lemma B.0.7.

We claim that $\mathrm{Orb}^{-1}(\pi)$ is $(2, n/2k)$-concentrated. To see this, let $\pi = \{P_1, \ldots, P_{n/k}\}$ and let $\mathrm{Sym}_k(P_i) \subset \mathrm{Sym}(n)$ be the set of all k-cycles with support in P_i. Then $\mathrm{Orb}^{-1}(\pi)$ is isometric to $\mathrm{Sym}_k(P_1) \times \cdots \times \mathrm{Sym}_k(P_{n/k})$. The diameter of $\mathrm{Sym}_k(P_i)$,

viewed as a subset of $\mathrm{Sym}(n)$ with the normalized Hamming metric on $\mathrm{Sym}(n)$, is k/n. So the claim follows from [32, Corollary 1.17]. This verifies condition (4) of Lemma B.0.7.

For $\pi_1, \pi_2 \in \mathrm{Part}(n, k; \chi, \vec{t})$, let X_{π_1,π_2} be the set of all pairs $(\sigma_1, \sigma_2) \in \mathrm{Orb}^{-1}(\pi_1) \times \mathrm{Orb}^{-1}(\pi_2)$ such that if $P \in \pi_1 \cap \pi_2$ then the restriction of σ_1 to P equals the restriction of σ_2 to P. Observe that X_{π_1,π_2} is non-empty and the projection maps $X_{\pi_1,\pi_2} \to \mathrm{Orb}^{-1}(\pi_i)$ $(i = 1, 2)$ are constant-to-1. In fact, for any $\sigma_1 \in \mathrm{Orb}^{-1}(\pi_1)$, the set of σ_2 with $(\sigma_1, \sigma_2) \in X_{\pi_1,\pi_2}$ is bijective with the set of assignments of k-cycles to parts in $\pi_2 \setminus \pi_1$.

Let μ_{π_1,π_2} be the uniform probability measure on X_{π_1,π_2}. Since the projection maps are constant-to-1, the marginals of μ_{π_1,π_2} are uniform. Moreover, if $(\sigma_1, \sigma_2) \in X_{\pi_1,\pi_2}$ then

$$\{i \in [n] : \ \sigma_1(i) \neq \sigma_2(i)\} \subset \cup_{P \in \pi_1 \setminus \pi_2} P.$$

Thus

$$d_{\mathrm{Sym}(n)}(\sigma_1, \sigma_2) \leq n^{-1} |\cup_{P \in \pi_1 \setminus \pi_2} P| = n^{-1} k |\pi_1 \vartriangle \pi_2|/2 = d(\pi_1, \pi_2).$$

This verifies condition (5) of Lemma B.0.7.

We have now verified all of the conditions of Lemma B.0.7. The lemma follows.

□

Let $\mathrm{Sym}(n, k; \chi)$ be the set of all $\sigma \in \mathrm{Sym}(n, k)$ such that if $\vec{t} = (t_j)_{j=0}^k$ is the type of σ with respect to χ then $t_0 = t_k = 0$. In other words, $\sigma \in \mathrm{Sym}(n, k; \chi)$ if and only if the orbit-partition π of σ is proper with respect to χ (where we think of π as a collection of hyper-edges).

Let $\vec{s} = (s_j)$ with $s_0 = s_k = 0$ and $s_j = \frac{1}{k(2^k-2)} \binom{k}{j}$ for $0 < j < k$. For $\delta > 0$ let $\mathrm{Sym}_\delta(n, k; \chi)$ be the set of all $\sigma \in \mathrm{Sym}(n, k; \chi)$ such that if $\vec{t} = (t_i)_{i=0}^k$ is the type of

σ (with respect to χ) then

$$\sum_{i=0}^{k} |s_i - t_i|^2 < \delta^2.$$

Lemma B.0.10. *With notation as above, for sufficiently large n*

$$\frac{|\mathrm{Sym}_\delta(n,k;\chi)|}{|\mathrm{Sym}(n,k;\chi)|} \geq 1 - e^{-\lambda_1 \delta^2 n}$$

where $\lambda_1 > 0$ is a constant depending only on k.

Proof. Let

$$\begin{aligned}
\mathrm{Part}(n,k;\chi,\vec{t}) &= \{\pi \in \mathrm{Part}(n,k) : \ \pi \text{ has type } \vec{t} \text{ with respect to } \chi\}, \\
\mathrm{Part}(n,k;\chi) &= \{\pi \in \mathrm{Part}(n,k) : \ \chi \text{ is proper with respect to } \pi\}, \\
\mathrm{Part}_\delta(n,k;\chi) &= \left\{\pi \in \mathrm{Part}(n,k;\chi) : \quad \text{if } \vec{t} \text{ is the type of } \pi \text{ with respect to } \chi \text{ then } \sum_{i=0}^{k} |t_i - s_i|^2 <
\end{aligned}$$

The orbit-partition map from $\mathrm{Sym}(n,k) \to \mathrm{Part}(n,k)$ is constant-to-1 and maps $\mathrm{Sym}(n,k;\chi)$ onto $\mathrm{Part}(n,k;\chi)$ and $\mathrm{Sym}_\delta(n,k;\chi)$ onto $\mathrm{Part}_\delta(n,k;\chi)$. Therefore, it suffices to prove

$$\frac{|\,\mathrm{Part}_\delta(n,k;\chi)|}{|\,\mathrm{Part}(n,k;\chi)|} \geq 1 - e^{-\lambda \delta^2 n}$$

where $\lambda > 0$ is a constant depending only on k.

Let $\widetilde{\mathcal{M}}$ be the set of all vectors $\vec{t} = (t_i)_{i=0}^k \in [0,1]^{k+1}$ such that $t_0 = t_k = 0$, $\sum_i t_i = 1/k$ and $\sum_j j t_j = 1/2$.

Recall from Lemma 3.3.8 that if $\vec{t} \in \widetilde{\mathcal{M}}$ and $n\vec{t}$ is $\mathbb{Z}$-valued then

$$(1/n)\log|\,\mathrm{Part}(n,k;\chi,\vec{t})| = (1-1/k)(\log(n)-1) - \log(2) + J(\vec{t}) + O(n^{-1}\log(n))$$

where $J(\vec{t}) = H(\vec{t}) - \sum_{j=0}^k t_j \log(j!(k-j)!)$. By the proof of Theorem 3.3.2 (specifically equation (3.10)), J is uniquely maximized in $\widetilde{\mathcal{M}}$ by the vector $\vec{s}$.

In order to get a lower bound on $|\,\mathrm{Part}(n,k;\chi)|$, observe that there exists $\vec{r} \in \widetilde{\mathcal{M}}$

such that $n\vec{r}$ is $\mathbb{Z}$-valued and $|s_i - r_i| \leq k/n$ for all i. Thus $J(\vec{r}) - J(\vec{s}) = O(1/n)$. It follows that

$$\frac{1}{n}\log|\operatorname{Part}(n,k;\chi)| \;\geq\; \frac{1}{n}\log|\operatorname{Part}(n,k;\chi,\vec{r})| = (1-1/k)(\log(n)-1) - \log(2) + J(\vec{s}) + O(n^{-1}\log$$

We claim that the Hessian of J is negative definite. To see this, one can consider J to be a function of $[0,1]^{k+1}$. The linear terms in J do not contribute to its Hessian. Since the second derivative of $x \mapsto -x\log x$ is $-1/x$,

$$\frac{\partial^2 J}{\partial t_i \partial t_j} = \begin{cases} 0 & i \neq j \\ -1/t_i & i = j \end{cases}$$

Thus the Hessian is diagonal and every eigenvalue is negative; so it is negative definite.

Thus if $\vec{t} \in \widetilde{\mathcal{M}}$ is such that $\sum_i |t_i - s_i|^2 \geq \delta^2$ then

$$(1/n)\log|\operatorname{Part}(n,k;\chi,\vec{t})| \leq (1-1/k)(\log(n)-1) - \log(2) + J(\vec{s}) - \delta^2\lambda_1' + O(n^{-1}\log(n))$$

where $\lambda_1' = \frac{1}{2}\min_{\vec{t}\in\widetilde{\mathcal{M}}}\min_{1\leq i\leq k-1} 1/t_i$ is half the smallest absolute value of an eigenvalue of the Hessian of J on $\widetilde{\mathcal{M}}$.

If $\vec{t}$ is the type of a k-partition π of n then $t_i \in \{0, 1/n, 2/n, \ldots, 1\}$. Thus the number of different types of k-partitions of $[n]$ is bounded by a polynomial in n (namely $(n+1)^{k+1}$). Thus

$$\begin{aligned} \frac{|\operatorname{Part}_\delta(n,k;\chi)|}{|\operatorname{Part}(n,k;\chi)|} &\geq 1 - (n+1)^{k+1}\frac{\exp(n[(1-1/k)(\log(n)-1) - \log(2) + J(\vec{s}) - \delta^2\lambda_1' + O(n^{-1}\log}{\exp(n[(1-1/k)(\log(n)-1) - \log(2) + J(\vec{s}) + O(n^{-1}\log(n)} \\ &= 1 - n^c\exp(-\delta^2\lambda_1' n) \end{aligned}$$

where $c = O_k(1)$. This implies the lemma.

□

Recall that a k**-cycle** is a permutation $\pi \in \mathrm{Sym}(n)$ of the form $\pi = (v_1, \ldots, v_k)$ for some $v_1, \ldots, v_k \in [n]$. In other words, π has $n - k$ fixed points and one orbit of size k. The **support** of $\pi \in \mathrm{Sym}(n)$ is the complement of the set of π-fixed points. It is denoted by $\mathrm{supp}(\pi)$. Two permutations are **disjoint** if their supports are disjoint. A permutation $\pi \in \mathrm{Sym}(n)$ is a **disjoint product of k-cycles** if there exist pairwise disjoint k-cycles $\pi_1, \ldots, \pi_m$ such that $\pi = \pi_1 \cdots \pi_m$. In this case we say that each π_i is **contained in** π.

Lemma B.0.11. *Let $\vec{t}, \vec{u} \in [0,1]^{k+1}$. Suppose*

$$\sum_{i=0}^{k} |t_i - u_i| < \delta.$$

Suppose $\mathrm{Sym}(n, k; \chi, \vec{t})$ and $\mathrm{Sym}(n, k; \chi, \vec{u})$ are non-empty (for some integer n and equitable coloring χ).

For $\sigma, \sigma' \in \mathrm{Sym}(n, k)$, let $|\sigma \triangle \sigma'|$ be the number of k-cycles τ that are either in σ or in σ' but not in both. Let

$$Z = \{(\sigma, \sigma') \in \mathrm{Sym}(n, k; \chi, \vec{t}) \times \mathrm{Sym}(n, k; \chi, \vec{u}) : \ |\sigma \triangle \sigma'| \leq \delta n\}.$$

Then Z is non-empty and there exists a probability measure μ on Z with marginals equal to the uniform probability measures on $\mathrm{Sym}(n, k; \chi, \vec{t})$ and $\mathrm{Sym}(n, k; \chi, \vec{u})$ respectively.

Proof. Let $\rho \in \mathrm{Sym}(n)$ be a disjoint product of k-cycles. The **type of ρ with respect to χ** is the vector $\vec{r} = (r_i)_{i=0}^{k}$ defined by: r_i is $1/n$ times the number of k-cycles ρ' contained in ρ such that $|\,\mathrm{supp}(\rho') \cap \chi^{-1}(1)| = i$.

Let $\sigma \in \mathrm{Sym}(n, k; \chi, \vec{t})$. Then there exist disjoint k-cycles $\sigma'_1, \ldots, \sigma'_m$ in σ such that if $\rho = \sigma'_1 \cdots \sigma'_m$ and $\vec{r} = (r_i)_{i=0}^{k}$ is the type of ρ then $r_i = \min(t_i, u_i)$. Note $m \geq n(1/k - \delta/2)$ by assumption on $\vec{t}$ and $\vec{u}$. Moreover, there exist k-cycles $\sigma'_{m+1}, \cdots, \sigma'_{n/k}$

such that the collection $\sigma'_1, \ldots, \sigma'_{n/k}$ is pairwise disjoint and the type of $\sigma' = \sigma'_1 \cdots \sigma'_{n/k}$ is $\vec{u}$. Then $|\sigma \triangle \sigma'| = 2(n/k - m) \leq \delta n$. So $(\sigma, \sigma') \in Z$ which proves Z is non-empty.

We claim that there is a constant $C_1 > 0$ such that for every $\sigma \in \mathrm{Sym}(n, k; \chi, \vec{t})$ the number of $\sigma' \in \mathrm{Sym}(n, k; \chi, \vec{u})$ with $(\sigma, \sigma') \in Z$ is C_1. Indeed the following algorithm constructs all such σ' with no duplications:

Step 1. Let $\sigma = \sigma_1 \cdots \sigma_{n/k}$ be a representation of σ as a disjoint product of k-cycles. Choose a vector $\vec{r} = (r_i)_{i=0}^{k}$ such that

1. there exists a subset $S \subset [n/k]$ with cardinality $|S| \geq n(1/k - \delta/2)$ such that if $\rho = \prod_{i \in S} \sigma_i$ then $\vec{r}$ is the type of ρ;

2. $r_i \leq u_i$ for all i.

Step 2. Choose a subset $S \subset [n/k]$ satisfying the condition in Step 1.

Step 3. Choose pairwise disjoint k-cycles $\sigma'_1, \ldots, \sigma'_{n/k-|S|}$ such that

1. $\mathrm{supp}(\sigma_i) \cap \mathrm{supp}(\sigma'_j) = \emptyset$ $(\forall i \in S)$ $(\forall j)$;

2. σ'_j is not contained in σ $(\forall j)$;

3. if $\sigma' = \prod_{i \in S} \sigma_i \prod_j \sigma'_j$ then σ' has type $\vec{u}$.

The range of possible vectors $\vec{r}$ in Step 1 depends only on $k, n, \vec{t}, \vec{u}$. The number of choices in Steps 2 and 3 depends only on the choice of $\vec{r}$ in Step 1 and on $k, n, \vec{t}, \vec{u}$. This proves the claim.

Similarly, there is a constant $C_2 > 0$ such that for every $\sigma' \in \mathrm{Sym}(n, k; \chi, \vec{u})$ the number of $\sigma \in \mathrm{Sym}(n, k; \chi, \vec{t})$ with $(\sigma, \sigma') \in Z$ is C_2. It follows that the uniform probability measure on Z has marginals equal to the uniform probability measures on $\mathrm{Sym}(n, k; \chi, \vec{t})$ and $\mathrm{Sym}(n, k; \chi, \vec{u})$ respectively.

□

Corollary B.0.12. *Let $U_{\mathrm{Sym}(n,k;\chi)}$ denote the uniform probability measure on $\mathrm{Sym}(n, k; \chi)$ and let $\mathbb{E}_{\mathrm{Sym}(n,k;\chi)}$ be the associated expectation operator. There are constants $c, \lambda > 0$*

(depending only on k) such that for every $\delta > 0$, there exists N_δ such that for all $n > N_\delta$, for every 1-Lipschitz $f : \mathrm{Sym}(n, k; \chi) \to \mathbb{R}$,

$$U_{\mathrm{Sym}(n,k;\chi)}\left(\left|f - \mathbb{E}_{\mathrm{Sym}(n,k;\chi)}[f]\right| > \delta\right) \leq c\exp(-\lambda\delta^2 n).$$

Moreover $\delta \mapsto N_\delta$ is monotone decreasing.

Proof. The set $\mathrm{Sym}(n, k; \chi)$ is the disjoint union of $\mathrm{Sym}(n, k; \chi, \vec{t})$ over $\vec{t} \in [0, 1]^{k+1}$. Let $\delta > 0$. Lemmas B.0.9, B.0.10 and B.0.11 imply that for all sufficiently large n, this decomposition of $\mathrm{Sym}(n, k; \chi)$ satisfies the criterion in Lemma B.0.5 where we set $c = 3$, $\eta = \exp(-\lambda_1\delta^2 n)$ and $\lambda = \lambda_0 n$ where $\lambda_0, \lambda_1 > 0$ depend only on k. So for every 1-Lipschitz function $f : \mathrm{Sym}(n, k; \chi) \to \mathbb{R}$, every $\epsilon > \delta + 2\eta$ and all sufficiently large n,

$$U_{\mathrm{Sym}(n,k;\chi)}\left(\left|f - \mathbb{E}_{\mathrm{Sym}(n,k;\chi)}[f]\right| > \epsilon\right) \leq \exp(-\lambda_1\delta^2 n) + c\exp\left(-\lambda_0 n\left(\epsilon - \delta - 2\eta\right)^2\right).$$

In particular, there exist N_δ such that if $n > N_\delta$ the inequality above holds and $2\eta < \delta$. By choosing N_δ larger if necessary, we require that $\delta \mapsto N_\delta$ is monotone decreasing.

Set $\epsilon = 3\delta$. Because $\epsilon - \delta - 2\eta \geq \delta$

$$U_{\mathrm{Sym}(n,k;\chi)}\left(\left|f - \mathbb{E}_{\mathrm{Sym}(n,k;\chi)}[f]\right| > 3\delta\right) \leq \exp(-\lambda_1\delta^2 n) + c\exp\left(-\lambda_0 n\delta^2\right) \leq (1 + c)\exp(-\lambda\delta^2 n)$$

where $\lambda = \min(\lambda_0, \lambda_1)$. The corollary is now finished by changing variables.

□

Proof of Theorem B.0.2. The space of homomorphisms $\mathrm{Hom}_\chi(\Gamma, \mathrm{Sym}(n))$ is the d-fold direct power of the spaces $\mathrm{Sym}(n, k; \chi)$. So the Theorem follows from Corollary B.0.12 and the proof of Lemma B.0.6.

□

Bibliography

[1] A. Coja-Oghlan, C. Efthymiou, and S. Hetterich. On the chromatic number of random regular graphs. *J. Combin. Theory Ser. B*, 116:367–439, 2016.

[2] J. Ding, A. Sly, and N. Sun. Maximum independent sets on random regular graphs. *Acta Math.*, 217(2):263–340, 2016.

[3] A. Dembo, A. Montanari, and S. Sen. Extremal cuts of sparse random graphs. *Ann. Probab.*, 45(2):1190–1217, 2017.

[4] J. Ding, A. Sly, and N. Sun. Proof of the satisfiability conjecture for large k [extended abstract]. In *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing*, pages 59–68. ACM, New York, 2015.

[5] E. Mossel, J. Neeman, and A. Sly. Consistency thresholds for the planted bisection model [extended abstract]. In *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing*, pages 69–75. ACM, New York, 2015.

[6] M. Mézard and A. Montanari. *Information, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[7] G. Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43:1754–1756, Dec 1979.

[8] D. Ruelle. A mathematical reformulation of derrida's rem and grem. *Comm. Math. Phys.*, 108:225–239, 1987.

[9] E. Bolthausen and A. Sznitman. On ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197:247–276, 1998.

[10] S. Ghirlanda and F. Guerra. General properties of overlap probability distributions in disordered systems. towards parisi ultrametricity. *J. Phys. A*, 31:9149–9155, 1998.

[11] M. Aizenman and P. Contucci. On the stability of the quenched state in mean-field spin-glass models. *J. Statist. Phys.*, 92:765–783, 1998.

[12] M. Talagrand. The parisi formula. *Annals of Mathematics*, 163(1):221–263, 2006.

[13] D. Panchenko. The parisi ultrametricity conjecture. *Ann. of Math.*, 177:383–393, 2013.

[14] D. Panchenko. *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer, New York, NY, 2013.

[15] A. Coja-Oghlan, C. Efthymiou, N. Jaafari, M. Kang, and T. Kapetanopoulos. Charting the replica symmetric phase. *Communications in Mathematical Physics*, 359(2):603 – 698, 2018.

[16] P. Carmona and Y. Hu. Universality in sherrington–kirkpatrick's spin glass model. *Annales de l'I.H.P. Probabilités et statistiques*, 42(2):215–222, 2006.

[17] D. Airey, L. Bowen, and F. Lin. A topological dynamical system with two different positive sofic entropies. *Trans. Amer. Math. Soc. Ser. B*, 9:35–98, 2022.

[18] R. L. Adler, A. G. Konheim, and M. H. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.*, 114:309–319, 1965.

[19] J. Moulin Ollagnier. *Ergodic theory and statistical mechanics*, volume 1115 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985.

[20] D. Kerr and H. Li. Entropy and the variational principle for actions of sofic groups. *Invent. Math.*, 186(3):501–558, 2011.

[21] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. *J. Amer. Math. Soc.*, 23(1):217–245, 2010.

[22] L. Bowen. Examples in the entropy theory of countable group actions. *Ergodic Theory Dynam. Systems*, 40(10):2593–2680, 2020.

[23] C. Bordenave. A new proof of friedman's second eigenvalue theorem and its extension to random lifts. *Annales scientifiques de l'École normale supérieure*, 53, 02 2015.

[24] D. Carroll. Periodic points in shifts of finite type over groups with connections to growth, 2016.

[25] C. Bordenave and B. Collins. Eigenvalues of random lifts and polynomials of random permutation matrices. *Ann. Math.*, 190:811–875, 2019.

[26] F. Lehner. Free operators with operator coefficients. *Colloquium Mathematicum*, 74(2):321–328, 1997.

[27] D. Achlioptas and C. Moore. Random k-SAT: two moments suffice to cross a sharp threshold. *SIAM J. Comput.*, 36(3):740–762, 2006.

[28] A. Coja-Oghlan and L. Zdeborova. The condensation transition in random hypergraph 2-coloring. *arXiv preprint arXiv:1107.2341*, 2011.

[29] A. Coja-Oghlan and L. Zdeborová. The condensation transition in random hypergraph 2-coloring. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 241–250. ACM, New York, 2012.

[30] T. Austin. Additivity properties of sofic entropy and measures on model spaces. *Forum Math. Sigma*, 4:e25, 79, 2016.

[31] D. Kerr and H. Li. Soficity, amenability, and dynamical entropy. *Amer. J. Math.*, 135(3):721–761, 2013.

[32] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.